

SEQUENCE LISTING



(1) GENERAL INFORMATION:

(i) APPLICANT: STUIVER, Maarten Hendrik  
CUSTERS, Jerome Humbertina Henricus Victor  
SELA-BURLAGE, Marianne Beatrix  
MELCHERS, Leo Sjoerd  
VAN DEVENTER-TROOST, Johanna Pieternella  
LAGEWEG, Wessel  
PONSTEIN, Anne Silene  
LAGEWEG, Wessel  
PONSTEIN, Anne Silene

(ii) TITLE OF INVENTION: ANTIFUNGAL PROTEINS, DNA CODING  
THEREFOR, AND HOSTS INCORPORATING  
SAME.

(iii) NUMBER OF SEQUENCES: 75

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: LADAS & PARRY  
(B) STREET: 26 WEST 61 STREET  
(C) CITY: NEW YORK  
(D) STATE: NY  
(E) COUNTRY: USA  
(F) ZIP: 10023 - 7604

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3.25" Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: WINDOWS 95  
(D) SOFTWARE: WORDPERFECT 8

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 09/258,031  
(B) FILING DATE: 25-FEB-1999  
(C) CLASSIFICATION: 435

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PCT/EP97/04923  
(B) FILING DATE: 04-SEP-1997

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: EP97200831.2  
(B) FILING DATE: 19-MAR-1997

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: EP96202466.7  
(B) FILING DATE: 04-SEP-1996

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helianthus annuus

(B) STRAIN: cv. zebulon

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Ser Ile Asn Val Asp Ile Glu Gln Glu Thr Ala Trp Val Gln Ala Gly  
1 5 10 15

Ala Thr Leu Gly Glu Val Tyr Tyr Arg  
20 25

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helianthus annuus

(B) STRAIN: cv. zebulon

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Asp Pro Ser Phe Pro Ile Thr Gly Glu Val Tyr Thr Pro Gly Xaa Ser  
1 5 10 15

Ser Phe Pro Thr Val Leu Gln Asn Tyr  
20 25

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: YES

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1  
(D) OTHER INFORMATION: /function= "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

AACTTCTCCN AGNGTNGCNC CNGCTTGNAC CCA

33

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: YES

(ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(B) LOCATION: 1  
(D) OTHER INFORMATION: /function= "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GATCCNTCTT TCCCNATTAC TGGNGAGGTT TA

32

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 354 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helianthus annuus  
(B) STRAIN: cv. zebulon

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 1..354

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GAT CCG TCT TTC CCG ATT ACT GGG GAG GTT TAC ACT CCC GGA AAC TCA  
Asp Pro Ser Phe Pro Ile Thr Gly Glu Val Tyr Thr Pro Gly Asn Ser  
1 5 10 15

48

TCT TTT CCT ACC GTC TTG CAA AAC TAC ATC CGA AAC CTT CGG TTC AAT	96
Ser Phe Pro Thr Val Leu Gln Asn Tyr Ile Arg Asn Leu Arg Phe Asn	
20 25 30	
GAA ACT ACC ACA CCA AAA CCC TTT TTA ATC ATC ACA GCC GAA CAT GTT	144
Glu Thr Thr Thr Pro Lys Pro Phe Leu Ile Ile Thr Ala Glu His Val	
35 40 45	
TCC CAC ATT CAG GCA GCT GTG GTT TGT GGC AAA CAA AAC CGG TTG CTA	192
Ser His Ile Gln Ala Ala Val Val Cys Gly Lys Gln Asn Arg Leu Leu	
50 55 60	
CTG AAA ACC AGA AGC GGT GGT CAT GAT TAT GAA GGT CTT TCC TAC CTT	240
Leu Lys Thr Arg Ser Gly Gly His Asp Tyr Glu Gly Leu Ser Tyr Leu	
65 70 75 80	
ACA AAC ACA AAC CAA CCC TTC TTC ATT GTG GAC ATG TTC AAT TTA AGG	288
Thr Asn Thr Asn Gln Pro Phe Phe Ile Val Asp Met Phe Asn Leu Arg	
85 90 95	
TCC ATA AAC GTA GAT ATC GAA CAA GAA ACC GCA TGG GTC CAA GCC GGC	336
Ser Ile Asn Val Asp Ile Glu Gln Glu Thr Ala Trp Val Gln Ala Gly	
100 105 110	
GCC ACC CTC GGA GAA GTT	354
Ala Thr Leu Gly Glu Val	
115	

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 118 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Asp Pro Ser Phe Pro Ile Thr Gly Glu Val Tyr Thr Pro Gly Asn Ser	
1 5 10 15	
Ser Phe Pro Thr Val Leu Gln Asn Tyr Ile Arg Asn Leu Arg Phe Asn	
20 25 30	
Glu Thr Thr Thr Pro Lys Pro Phe Leu Ile Ile Thr Ala Glu His Val	
35 40 45	
Ser His Ile Gln Ala Ala Val Val Cys Gly Lys Gln Asn Arg Leu Leu	
50 55 60	
Leu Lys Thr Arg Ser Gly Gly His Asp Tyr Glu Gly Leu Ser Tyr Leu	
65 70 75 80	
Thr Asn Thr Asn Gln Pro Phe Phe Ile Val Asp Met Phe Asn Leu Arg	
85 90 95	
Ser Ile Asn Val Asp Ile Glu Gln Glu Thr Ala Trp Val Gln Ala Gly	
100 105 110	

Ala Thr Leu Gly Glu Val  
115

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1
- (D) OTHER INFORMATION: /function= "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CAGGCAGCTG TGGTTTGTGG C

21

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1
- (D) OTHER INFORMATION: /function= "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GTCCACAATG AAGAAGGGTT G

21

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1

(D) OTHER INFORMATION: /function= "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

ACGTAGATAT CGAACAAGAA ACCGC

25

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GCTTTACTAC ACGGGCTTCC CCAG

24

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CTGGGGAAGC CCGTGTAGTA AAGC

24

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GGTACTCCAA CCACGGCGCT C

21

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

CGGGAAGTTG CAGAAGATTG GGTG

25

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GAGCAAGAGA AGAAGGAGAC

20

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1784 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helianthus annuus*

(B) STRAIN: Zebulon

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 21..1608

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

ATATCACATC TTCTTTCAAC ATG CAA ACT TCC ATT CTT ACT CTC CTT CTT	50
Met Gln Thr Ser Ile Leu Thr Leu Leu Leu	
1 5 10	
CTC TTG CTC TCA ACC CAA TCT TCT GCA ACT TCC CGT TCC ATT ACA GAT	98
Leu Leu Leu Ser Thr Gln Ser Ser Ala Thr Ser Arg Ser Ile Thr Asp	
15 20 25	
CGC TTC ATT CAA TGT TTA CAC GAC CGG GCC GAC CCT TCA TTT CCG ATA	146
Arg Phe Ile Gln Cys Leu His Asp Arg Ala Asp Pro Ser Phe Pro Ile	
30 35 40	
ACC GGA GAG GTT TAC ACT CCC GGA AAC TCA TCT TTT CCT ACC GTC TTG	194
Thr Gly Glu Val Tyr Thr Pro Gly Asn Ser Ser Phe Pro Thr Val Leu	
45 50 55	
CAA AAC TAC ATC CGA AAC CTT CGG TTC AAT GAA ACT ACC ACA CCA AAA	242
Gln Asn Tyr Ile Arg Asn Leu Arg Phe Asn Glu Thr Thr Thr Pro Lys	
60 65 70	
CCC TTT TTA ATC ATC ACA GCC GAA CAT GTT TCC CAC ATT CAG GCA GCT	290
Pro Phe Leu Ile Ile Thr Ala Glu His Val Ser His Ile Gln Ala Ala	
75 80 85 90	
GTG GTT TGT GGC AAA CAA AAC CGG TTG CTA CTG AAA ACC AGA AGC GGT	338
Val Val Cys Gly Lys Gln Asn Arg Leu Leu Lys Thr Arg Ser Gly	
95 100 105	
GGT CAT GAT TAT GAA GGT CTT TCC TAC CTT ACA AAC ACA AAC CAA CCC	386
Gly His Asp Tyr Glu Gly Leu Ser Tyr Leu Thr Asn Thr Asn Gln Pro	
110 115 120	
TTC TTC ATT GTG GAC ATG TTC AAT TTA AGG TCC ATA AAC GTA GAT ATC	434
Phe Phe Ile Val Asp Met Phe Asn Leu Arg Ser Ile Asn Val Asp Ile	
125 130 135	
GAA CAA GAA ACC GCA TGG GTC CAA GCC GGT GCG ACT CTT GGT GAA GTG	482
Glu Gln Glu Thr Ala Trp Val Gln Ala Gly Ala Thr Leu Gly Glu Val	
140 145 150	
TAC TAT CGA ATA GCG GAG AAA AGT AAC AAG CAT GGT TTT CCG GCA GGG	530
Tyr Tyr Arg Ile Ala Glu Lys Ser Asn Lys His Gly Phe Pro Ala Gly	
155 160 165 170	
GTT TGT CCA ACG GTT GGC GTT GGT GGG CAT TTT AGT GGT GGT GGG TAT	578
Val Cys Pro Thr Val Gly Val Gly Gly His Phe Ser Gly Gly Gly Tyr	
175 180 185	
GGT AAT TTG ATG AGA AAA TAT GGT TTG TCG GTT GAT AAT ATT GTT GAT	626
Gly Asn Leu Met Arg Lys Tyr Gly Leu Ser Val Asp Asn Ile Val Asp	
190 195 200	



GCT CAA ATA ATA GAT GTG AAT GGC AAG CTT TTG GAT CGA AAG AGT ATG	674
Ala Gln Ile Ile Asp Val Asn Gly Lys Leu Leu Asp Arg Lys Ser Met	
205 210 215	
GGT GAG GAT TTG TTT TGG GCG ATC ACC GGC GGT GGT GGT GTT AGT TTT	722
Gly Glu Asp Leu Phe Trp Ala Ile Thr Gly Gly Gly Gly Val Ser Phe	
220 225 230	
GGT GTG GTT CTA GCC TAC AAA ATC AAA CTA GTT CGT GTT CCG GAG GTT	770
Gly Val Val Leu Ala Tyr Lys Ile Lys Leu Val Arg Val Pro Glu Val	
235 240 245 250	
GTG ACC GTG TTT ACC ATT GAA AGA AGA GAG GAA CAA AAC CTC AGC ACC	818
Val Thr Val Phe Thr Ile Glu Arg Arg Glu Glu Gln Asn Leu Ser Thr	
255 260 265	
ATC GCG GAA CGA TGG GTA CAA GTT GCT GAT AAG CTA GAT AGA GAT CTT	866
Ile Ala Glu Arg Trp Val Gln Val Ala Asp Lys Leu Asp Arg Asp Leu	
270 275 280	
TTC CTT CGA ATG ACC TTT AGT GTC ATA AAC GAT ACC AAC GGT GGA AAG	914
Phe Leu Arg Met Thr Phe Ser Val Ile Asn Asp Thr Asn Gly Gly Lys	
285 290 295	
ACA GTC CGT GCT ATC TTT CCA ACG TTG TAC CTT GGA AAC TCG AGG AAT	962
Thr Val Arg Ala Ile Phe Pro Thr Leu Tyr Leu Gly Asn Ser Arg Asn	
300 305 310	
CTT GTT ACA CTT TTG AAT AAA GAT TTC CCC GAG TTA GGG TTG CAA GAA	1010
Leu Val Thr Leu Leu Asn Lys Asp Phe Pro Glu Leu Gly Leu Gln Glu	
315 320 325 330	
TCG GAT TGT ACT GAA ATG AGT TGG GTT GAG TCT GTG CTT TAC TAC ACG	1058
Ser Asp Cys Thr Glu Met Ser Trp Val Glu Ser Val Leu Tyr Tyr Thr	
335 340 345	
GGC TTC CCC AGT GGT ACT CCA ACC ACG GCG CTC TTA AGC CGT ACT CCT	1106
Gly Phe Pro Ser Gly Thr Pro Thr Thr Ala Leu Leu Ser Arg Thr Pro	
350 355 360	
CAA AGA CTC AAC CCA TTC AAG ATC AAA TCC GAT TAT GTG CAA AAT CCT	1154
Gln Arg Leu Asn Pro Phe Lys Ile Lys Ser Asp Tyr Val Gln Asn Pro	
365 370 375	
ATT TCT AAA CGA CAG TTC GAG TTC ATC TTC GAA AGG CTG AAA GAA CTT	1202
Ile Ser Lys Arg Gln Phe Glu Phe Ile Phe Glu Arg Leu Lys Glu Leu	
380 385 390	
GAA AAC CAA ATG TTG GCT TTC AAC CCA TAT GGT GGT AGA ATG AGT GAA	1250
Glu Asn Gln Met Leu Ala Phe Asn Pro Tyr Gly Gly Arg Met Ser Glu	
395 400 405 410	
ATA TCC GAA TTC GCA AAG CCT TTC CCA CAT AGA TCG GGT AAC ATA GCG	1298
Ile Ser Glu Phe Ala Lys Pro Phe Pro His Arg Ser Gly Asn Ile Ala	
415 420 425	
AAA ATT CAA TAC GAA GTA AAC TGG GAG GAT CTT AGC GAT GAA GCC GAA	1346
Lys Ile Gln Tyr Glu Val Asn Trp Glu Asp Leu Ser Asp Glu Ala Glu	
430 435 440	

AAT CGT TAC TTG AAT TTC ACA AGG CTG ATG TAT GAT TAC ATG ACC CCA 1394  
Asn Arg Tyr Leu Asn Phe Thr Arg Leu Met Tyr Asp Tyr Met Thr Pro  
445 450 455  
  
TTT GTG TCG AAA AAC CCT AGA AAA GCA TTT TTG AAC TAT AGG GAT TTG 1442  
Phe Val Ser Lys Asn Pro Arg Lys Ala Phe Leu Asn Tyr Arg Asp Leu  
460 465 470  
  
GAT ATT GGT ATC AAC AGC CAT GGC AGG AAT GCT TAT ACT GAA GGA ATG 1490  
Asp Ile Gly Ile Asn Ser His Gly Arg Asn Ala Tyr Thr Glu Gly Met  
475 480 485 490  
  
GTT TAT GGG CAC AAG TAT TTC AAA GAG ACA AAT TAC AAG AGG CTA GTA 1538  
Val Tyr Gly His Lys Tyr Phe Lys Glu Thr Asn Tyr Lys Arg Leu Val  
495 500 505  
  
AGT GTG AAG ACT AAA GTT GAT CCT GAC AAC TTC TTT AGG AAT GAG CAA 1586  
Ser Val Lys Thr Lys Val Asp Pro Asp Asn Phe Phe Arg Asn Glu Gln  
510 515 520  
  
AGC ATC CCA ACT TTG TCA TCT T GAAGAACGTA CATATATAAAA TAAATACCTT 1638  
Ser Ile Pro Thr Leu Ser Ser  
525  
  
TGTGCATGGT ATTTTCAGGG TGTTAAAGTG ATATTCAGAT ATTTATGATA GAATTTTGAC 1698  
  
TTGTATTTTA TACAATCAAA ATTGTATGGT TCTCCGAATT TCTCTTTTTA ATTCTGAAAA 1758  
  
ATACATATTA GTATTGTCAA AAAAAA 1784

(2) INFORMATION FOR SEQ ID NO: 16:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 529 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Met Gln Thr Ser Ile Leu Thr Leu Leu Leu Leu Leu Ser Thr Gln  
1 5 10 15  
Ser Ser Ala Thr Ser Arg Ser Ile Thr Asp Arg Phe Ile Gln Cys Leu  
20 25 30  
His Asp Arg Ala Asp Pro Ser Phe Pro Ile Thr Gly Glu Val Tyr Thr  
35 40 45  
Pro Gly Asn Ser Ser Phe Pro Thr Val Leu Gln Asn Tyr Ile Arg Asn  
50 55 60  
Leu Arg Phe Asn Glu Thr Thr Thr Pro Lys Pro Phe Leu Ile Ile Thr  
65 70 75 80  
Ala Glu His Val Ser His Ile Gln Ala Ala Val Val Cys Gly Lys Gln  
85 90 95

Asn Arg Leu Leu Leu Lys Thr Arg Ser Gly Gly His Asp Tyr Glu Gly  
 100 105 110  
 Leu Ser Tyr Leu Thr Asn Thr Asn Gln Pro Phe Phe Ile Val Asp Met  
 115 120 125  
 Phe Asn Leu Arg Ser Ile Asn Val Asp Ile Glu Gln Glu Thr Ala Trp  
 130 135 140  
 Val Gln Ala Gly Ala Thr Leu Gly Glu Val Tyr Tyr Arg Ile Ala Glu  
 145 150 155 160  
 Lys Ser Asn Lys His Gly Phe Pro Ala Gly Val Cys Pro Thr Val Gly  
 165 170 175  
 Val Gly Gly His Phe Ser Gly Gly Gly Tyr Gly Asn Leu Met Arg Lys  
 180 185 190  
 Tyr Gly Leu Ser Val Asp Asn Ile Val Asp Ala Gln Ile Ile Asp Val  
 195 200 205  
 Asn Gly Lys Leu Leu Asp Arg Lys Ser Met Gly Glu Asp Leu Phe Trp  
 210 215 220  
 Ala Ile Thr Gly Gly Gly Gly Val Ser Phe Gly Val Val Leu Ala Tyr  
 225 230 235 240  
 Lys Ile Lys Leu Val Arg Val Pro Glu Val Val Thr Val Phe Thr Ile  
 245 250 255  
 Glu Arg Arg Glu Glu Gln Asn Leu Ser Thr Ile Ala Glu Arg Trp Val  
 260 265 270  
 Gln Val Ala Asp Lys Leu Asp Arg Asp Leu Phe Leu Arg Met Thr Phe  
 275 280 285  
 Ser Val Ile Asn Asp Thr Asn Gly Gly Lys Thr Val Arg Ala Ile Phe  
 290 295 300  
 Pro Thr Leu Tyr Leu Gly Asn Ser Arg Asn Leu Val Thr Leu Leu Asn  
 305 310 315 320  
 Lys Asp Phe Pro Glu Leu Gly Leu Gln Glu Ser Asp Cys Thr Glu Met  
 325 330 335  
 Ser Trp Val Glu Ser Val Leu Tyr Tyr Thr Gly Phe Pro Ser Gly Thr  
 340 345 350  
 Pro Thr Thr Ala Leu Leu Ser Arg Thr Pro Gln Arg Leu Asn Pro Phe  
 355 360 365  
 Lys Ile Lys Ser Asp Tyr Val Gln Asn Pro Ile Ser Lys Arg Gln Phe  
 370 375 380  
 Glu Phe Ile Phe Glu Arg Leu Lys Glu Leu Glu Asn Gln Met Leu Ala  
 385 390 395 400  
 Phe Asn Pro Tyr Gly Gly Arg Met Ser Glu Ile Ser Glu Phe Ala Lys  
 405 410 415

Pro Phe Pro His Arg Ser Gly Asn Ile Ala Lys Ile Gln Tyr Glu Val  
                   420                                  425                                  430  
 Asn Trp Glu Asp Leu Ser Asp Glu Ala Glu Asn Arg Tyr Leu Asn Phe  
                   435                                  440                                  445  
 Thr Arg Leu Met Tyr Asp Tyr Met Thr Pro Phe Val Ser Lys Asn Pro  
                   450                                  455                                  460  
 Arg Lys Ala Phe Leu Asn Tyr Arg Asp Leu Asp Ile Gly Ile Asn Ser  
                   465                                  470                                  475                                  480  
 His Gly Arg Asn Ala Tyr Thr Glu Gly Met Val Tyr Gly His Lys Tyr  
                                   485                                  490                                  495  
 Phe Lys Glu Thr Asn Tyr Lys Arg Leu Val Ser Val Lys Thr Lys Val  
                                   500                                  505                                  510  
 Asp Pro Asp Asn Phe Phe Arg Asn Glu Gln Ser Ile Pro Thr Leu Ser  
                   515                                  520                                  525  
 Ser

(2) INFORMATION FOR SEQ ID NO: 17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

CCGCCATGGA GACTTCCATT CTTACTC

27

(2) INFORMATION FOR SEQ ID NO: 18:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 33 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GCCGGATCCT CAAGATGACA AAGTTGGGAT GCT

33

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1589 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helianthus annuus*
- (B) STRAIN: Zebulon

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1590

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

ATG GAG ACT TCC ATT CTT ACT CTC CTT CTT CTC TTG CTC TCA ACC CAA	48
Met Glu Thr Ser Ile Leu Thr Leu Leu Leu Leu Leu Ser Thr Gln	
1 5 10 15	
TCT TCT GCA ACT TCC CGT TCC ATT ACA GAT CGC TTC ATT CAA TGT TTA	96
Ser Ser Ala Thr Ser Arg Ser Ile Thr Asp Arg Phe Ile Gln Cys Leu	
20 25 30	
CAC GAC CGG GCC GAC CCT TCA TTT CCG ATA ACC GGA GAG GTT TAC ACT	144
His Asp Arg Ala Asp Pro Ser Phe Pro Ile Thr Gly Glu Val Tyr Thr	
35 40 45	
CCC GGA AAC TCA TCT TTT CCT ACC GTC TTG CAA AAC TAC ATC CGA AAC	192
Pro Gly Asn Ser Ser Phe Pro Thr Val Leu Gln Asn Tyr Ile Arg Asn	
50 55 60	
CTT CGG TTC AAT GAA ACT ACC ACA CCA AAA CCC TTT TTA ATC ATC ACA	240
Leu Arg Phe Asn Glu Thr Thr Thr Pro Lys Pro Phe Leu Ile Ile Thr	
65 70 75 80	
GCC GAA CAT GTT TCC CAC ATT CAG GCA GCT GTG GTT TGT GGC AAA CAA	288
Ala Glu His Val Ser His Ile Gln Ala Ala Val Val Cys Gly Lys Gln	
85 90 95	
AAC CGG TTG CTA CTG AAA ACC AGA AGC GGT GGT CAT GAT TAT GAA GGT	336
Asn Arg Leu Leu Leu Lys Thr Arg Ser Gly Gly His Asp Tyr Glu Gly	
100 105 110	
CTT TCC TAC CTT ACA AAC ACA AAC CAA CCC TTC TTC ATT GTG GAC ATG	384
Leu Ser Tyr Leu Thr Asn Thr Asn Gln Pro Phe Phe Ile Val Asp Met	
115 120 125	
TTC AAT TTA AGG TCC ATA AAC GTA GAT ATC GAA CAA GAA ACC GCA TGG	432
Phe Asn Leu Arg Ser Ile Asn Val Asp Ile Glu Gln Glu Thr Ala Trp	
130 135 140	

GTC CAA GCC GGT GCG ACT CTT GGT GAA GTG TAC TAT CGA ATA GCG GAG	480
Val Gln Ala Gly Ala Thr Leu Gly Glu Val Tyr Tyr Arg Ile Ala Glu	
145 150 155 160	
AAA AGT AAC AAG CAT GGT TTT CCG GCA GGG GTT TGT CCA ACG GTT GGC	528
Lys Ser Asn Lys His Gly Phe Pro Ala Gly Val Cys Pro Thr Val Gly	
165 170 175	
GTT GGT GGG CAT TTT AGT GGT GGT GGG TAT GGT AAT TTG ATG AGA AAA	576
Val Gly Gly His Phe Ser Gly Gly Gly Tyr Gly Asn Leu Met Arg Lys	
180 185 190	
TAT GGT TTG TCG GTT GAT AAT ATT GTT GAT GCT CAA ATA ATA GAT GTG	624
Tyr Gly Leu Ser Val Asp Asn Ile Val Asp Ala Gln Ile Ile Asp Val	
195 200 205	
AAT GGC AAG CTT TTG GAT CGA AAG AGT ATG GGT GAG GAT TTG TTT TGG	672
Asn Gly Lys Leu Leu Asp Arg Lys Ser Met Gly Glu Asp Leu Phe Trp	
210 215 220	
GCG ATC ACC GGC GGT GGT GGT GTT AGT TTT GGT GTG GTT CTA GCC TAC	720
Ala Ile Thr Gly Gly Gly Gly Val Ser Phe Gly Val Val Leu Ala Tyr	
225 230 235 240	
AAA ATC AAA CTA GTT CGT GTT CCG GAG GTT GTG ACC GTG TTT ACC ATT	768
Lys Ile Lys Leu Val Arg Val Pro Glu Val Val Thr Val Phe Thr Ile	
245 250 255	
GAA AGA AGA GAG GAA CAA AAC CTC AGC ACC ATC GCG GAA CGA TGG GTA	816
Glu Arg Arg Glu Glu Gln Asn Leu Ser Thr Ile Ala Glu Arg Trp Val	
260 265 270	
CAA GTT GCT GAT AAG CTA GAT AGA GAT CTT TTC CTT CGA ATG ACC TTT	864
Gln Val Ala Asp Lys Leu Asp Arg Asp Leu Phe Leu Arg Met Thr Phe	
275 280 285	
AGT GTC ATA AAC GAT ACC AAC GGT GGA AAG ACA GTC CGT GCT ATC TTT	912
Ser Val Ile Asn Asp Thr Asn Gly Gly Lys Thr Val Arg Ala Ile Phe	
290 295 300	
CCA ACG TTG TAC CTT GGA AAC TCG AGG AAT CTT GTT ACA CTT TTG AAT	960
Pro Thr Leu Tyr Leu Gly Asn Ser Arg Asn Leu Val Thr Leu Leu Asn	
305 310 315 320	
AAA GAT TTC CCC GAG TTA GGG TTG CAA GAA TCG GAT TGT ACT GAA ATG	1008
Lys Asp Phe Pro Glu Leu Gly Leu Gln Glu Ser Asp Cys Thr Glu Met	
325 330 335	
AGT TGG GTT GAG TCT GTG CTT TAC TAC ACG GGC TTC CCC AGT GGT ACT	1056
Ser Trp Val Glu Ser Val Leu Tyr Tyr Thr Gly Phe Pro Ser Gly Thr	
340 345 350	
CCA ACC ACG GCG CTC TTA AGC CGT ACT CCT CAA AGA CTC AAC CCA TTC	1104
Pro Thr Thr Ala Leu Leu Ser Arg Thr Pro Gln Arg Leu Asn Pro Phe	
355 360 365	
AAG ATC AAA TCC GAT TAT GTG CAA AAT CCT ATT TCT AAA CGA CAG TTC	1152
Lys Ile Lys Ser Asp Tyr Val Gln Asn Pro Ile Ser Lys Arg Gln Phe	
370 375 380	

[illegible]

Pro Gly Asn Ser Ser Phe Pro Thr Val Leu Gln Asn Tyr Ile Arg Asn  
 50 55 60

Leu Arg Phe Asn Glu Thr Thr Thr Pro Lys Pro Phe Leu Ile Ile Thr  
 65 70 75 80

Ala Glu His Val Ser His Ile Gln Ala Ala Val Val Cys Gly Lys Gln  
 85 90 95

Asn Arg Leu Leu Leu Lys Thr Arg Ser Gly Gly His Asp Tyr Glu Gly  
 100 105 110

Leu Ser Tyr Leu Thr Asn Thr Asn Gln Pro Phe Phe Ile Val Asp Met  
 115 120 125

Phe Asn Leu Arg Ser Ile Asn Val Asp Ile Glu Gln Glu Thr Ala Trp  
 130 135 140

Val Gln Ala Gly Ala Thr Leu Gly Glu Val Tyr Tyr Arg Ile Ala Glu  
 145 150 155 160

Lys Ser Asn Lys His Gly Phe Pro Ala Gly Val Cys Pro Thr Val Gly  
 165 170 175

Val Gly Gly His Phe Ser Gly Gly Gly Tyr Gly Asn Leu Met Arg Lys  
 180 185 190

Tyr Gly Leu Ser Val Asp Asn Ile Val Asp Ala Gln Ile Ile Asp Val  
 195 200 205

Asn Gly Lys Leu Leu Asp Arg Lys Ser Met Gly Glu Asp Leu Phe Trp  
 210 215 220

Ala Ile Thr Gly Gly Gly Gly Val Ser Phe Gly Val Val Leu Ala Tyr  
 225 230 235 240

Lys Ile Lys Leu Val Arg Val Pro Glu Val Val Thr Val Phe Thr Ile  
 245 250 255

Glu Arg Arg Glu Glu Gln Asn Leu Ser Thr Ile Ala Glu Arg Trp Val  
 260 265 270

Gln Val Ala Asp Lys Leu Asp Arg Asp Leu Phe Leu Arg Met Thr Phe  
 275 280 285

Ser Val Ile Asn Asp Thr Asn Gly Gly Lys Thr Val Arg Ala Ile Phe  
 290 295 300

Pro Thr Leu Tyr Leu Gly Asn Ser Arg Asn Leu Val Thr Leu Leu Asn  
 305 310 315 320

Lys Asp Phe Pro Glu Leu Gly Leu Gln Glu Ser Asp Cys Thr Glu Met  
 325 330 335

Ser Trp Val Glu Ser Val Leu Tyr Tyr Thr Gly Phe Pro Ser Gly Thr  
 340 345 350

Pro Thr Thr Ala Leu Leu Ser Arg Thr Pro Gln Arg Leu Asn Pro Phe  
 355 360 365



Lys Ile Lys Ser Asp Tyr Val Gln Asn Pro Ile Ser Lys Arg Gln Phe  
 370 375 380  
 Glu Phe Ile Phe Glu Arg Met Lys Glu Leu Glu Asn Gln Met Leu Ala  
 385 390 395 400  
 Phe Asn Pro Tyr Gly Gly Arg Met Ser Glu Ile Ser Glu Phe Ala Lys  
 405 410 415  
 Pro Phe Pro His Arg Ser Gly Asn Ile Ala Lys Ile Gln Tyr Glu Val  
 420 425 430  
 Asn Trp Glu Asp Leu Ser Asp Glu Ala Glu Asn Arg Tyr Leu Asn Phe  
 435 440 445  
 Thr Arg Leu Met Tyr Asp Tyr Met Thr Pro Phe Val Ser Lys Asn Pro  
 450 455 460  
 Arg Glu Ala Phe Leu Asn Tyr Arg Asp Leu Asp Ile Gly Ile Asn Ser  
 465 470 475 480  
 His Gly Arg Asn Ala Tyr Thr Glu Gly Met Val Tyr Gly His Lys Tyr  
 485 490 495  
 Phe Lys Glu Thr Asn Tyr Lys Arg Leu Val Ser Val Lys Thr Lys Val  
 500 505 510  
 Asp Pro Asp Asn Phe Phe Arg Asn Glu Gln Ser Ile Pro Thr Leu Ser  
 515 520 525  
 Ser

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 350 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..350

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GAGAAACTCG GAGACTTTCA CACAATGCCT AACCTCAAAC TCCGACCCCA AACATCCCAT 60  
 CTCCCCGCT ATCTTCTTCT CCGGAAATGG CTCCTACTCC TCCGTATTAC AAGCCAACAT 120  
 CCGTAACCTC CGCTTCAACA CCACCTCAAC TCCGAAACCC TTCCTCATAA TCGCCGCAAC 180

ACATGAATCC CATGTGCAAG CCGCGATTAC TTGCGGGAAA CGCCACAACC TTCAGATGAA 240  
 AATCAGAAGT GGAGGCCACG ACTACGATGG CTTGTCATAC GTTACATACT CTGGCAAACC 300  
 GTTCTTCGTC CTCGACATGT TTAACCTCCG TTCGGTGGAT GTCGACGTGG 350

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 278 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..278

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

GGCATGGATC TCCGCCGGAG CGACTCTCGG AGAGGTTTAT TATCGGATTT GGGAGAAAAG 60  
 CAGAGTCCAT GGATTCCCCG CCGGAGTTTG ACCGACGGTT GGTGTTGGTG GGCATTTAAG 120  
 CGGCGGTGGT TACGGTAACA TGGTGAGGAA GTTTGGATTA TCTGTGGATT ACGTTGAGGA 180  
 TGCCAAGATC GTCGATGTAA ACNGTCGGGT TTTAGATCGG AAAGCAATGG GTGAGGATCT 240  
 GTTCTGGGCG ATTACCGGTG GAGGAGGAGG TAGCGTAC 278

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 345 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 2..345

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

TGGACATATT AGCGGAGGAG GATTCGGTAC AATAATGAGG AAATACGGTT TAGCGTCTGA 60  
TAACGTTGTG GACGCACGTT TGATGGATGT AAATGGGAAA ACTCTTGACC GGAAAACGAT 120  
GGGAGAGGAT TTGTTTTGGG CGCTTAGAGG CCGTGGAGCT GCGAGTTTTG GCGTTGTCTT 180  
GTCGTGGAAG GTTAAGCTTG CTAGGGTTCC TGAAAAGGTA ACTTGTTTCA TAAGTCAACA 240  
TCCGATGGGA CCTAGCATGA ACAAGCTTGT TCATAGATGG CAATCCATAG GATCAAGANN 300  
GCTAGACGAA GATTTATTCA TCAGAGTCAA TATTGACAAC AGTCT 345

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 695 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana  
(B) STRAIN: ecotype Columbia

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 1..695

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

GTTTCGTTAAA ACCTATCCTN NANGGGCNAAG AGNATATCAA AGNTTGNTTA NGNAACCCAA 60  
NATTTCTGAA CTGGCCNCCT TCGGTGGTAT ATGNCNAAAN CCCTTGAATC TGCGNANCCN 120  
ATTCCGCATA GAAACGGAAC CCTCTTCAAG ATTCTCTATT TACNCGAACT GNCTAGANNG 180  
AATGACAAGA CATCGAGTAG NAAAATCAAC TGGATCAAAG AGATATACAA TTACATGGCG 240  
CCTTATGTCT CAAGCAATCC AAGACAAGCA TATGTGAACT ACAGAGATCT AGACTTCGGA 300  
CAGAACAAGA ACAACGCAAA GGTAACTTC ATTGAAGCTA AAATCTGGGG ACCTAAGTAC 360  
TTCAAAGGCA ATTTTGACAG ATTGGTGAAG ATTTAAACCA AGGTTGATCC AGAGAACTTC 420  
TTCAGGCACG AGCAGAGTAT CCCACCTATG CCCTACTAGA AGCTAGGTTC ATGAAACCAA 480  
TAACATTATC AAAAATAAGR ATAAATGRTA ATTGTATACA ACATGATTCG KCTTTCTTTA 540  
TTTCAGACAA TGTGGACACT ACTCTAAANT AAAAWGTCNA TTTACCTTAA AAAAAAATA 600

ATCCCCNNTA ANANAAAANT GGGGGGGCCN TTTTGGGGN TCCCGGTTTT NGGACGGGGN 660  
GCTTTNGGGG GGCTTGGNNT TTTTNGGN GCCCC 695

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 495 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..495

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

TCTGTTTTNA GGCAGAGCAG AGGAAGTTGT TGCTTGCTT GGTAAGGAGT TTCCTGAATT 60  
NAGTTTAAAG AAGGAGAACT GTTCGGAGAT GACTTGTTTT CAGTCAGCTT TATGGTGGGA 120  
TAATCGTGTT AACCTACTC ANATTGATCC WAAAGTGTTC CTCGATCGGA ATCTTGATAG 180  
AGCGAATTTT GGAAAGAGGA AATCGGATTA CGTTGCGAGT AAGATTCCTA GAGATGGGAT 240  
TAAGYCTTTT TCCAAGARGA TGMCTGACCT GGGGAAAAYC GGGCTTGTTT TTAAWCCGTA 300  
TGGTGGGAAA ATGGCGGAGG TTACGGTTAA CGCGACGCCG TTTCNCACC GAAGCAAGCT 360  
TTTTAAGATT CAGTACTCGG TGACTIONGCA AGAAAACCTCT NTCGAGATAG AGAAAGGGTT 420  
TCTTGAATCA GGCTAACGTC CTTATAGGTT CATGACCGGG TTTTNAGCA AGANCCCTGG 480  
AATNCTTACT TNAAT 495

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 204 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..204

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

AAATTAAAAC AAATCAATGT TGATATTGAA TCCAATAGTG CTTGGTTTCA ACCTGGTGCT 60  
ACGCTTGGTG AGCTTTACTA CAGAATTNCA GAGAAGAGCA AAATCCATGG ATTTCCNGCG 120  
GGTTTNTNCA CAAGCNTAGG CATAGGTGGG TATATNANAG GCGGTGGATA CGGTACCTTG 180  
ATGAGGAAGT ATGGTCTTNC GGGA 204

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 491 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..491

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

GAGATTTCTC GAGCAAGATA CTCCACTGAT GATCTTTGAG CCATTGGGTG GGAAAATCAG 60  
CAAGATTTCA GAAACAGAAT CTCCATATCC ACACAGAAGA GGTAATCTGT ATAATATACA 120  
GTACATGGTG AAATGGAAAAG TGAATGANGT CGAGGAGATG AACAAACATG TCAGGTGGAT 180  
GAGATCGTTA CACGATTACA TGACTCCGTA TGTTTCTAAA TCGCCGAGAG GAGCTTATTT 240  
GANTTACAGA GATCTTGATT TGGGCTCGAC CAAAGGGATT AACACGGGTT TCGGAGATGC 300  
AAGGAAATGG NNGGGTGAGN CTTTTTTCAA AGGTAATTTT CAAGGGGTTA GGTTTTGGTT 360  
AAAGGGGAGG TTTNNCCCAN CAAATTTTTT TTCAGGANCC GGCCANGNTT TTCCCCCCCC 420  
TNTTTTNGG NCCCCAATCN AAANCCCCGT TTAAAAGGG GGGCCATTTT NTTTTTTNCA 480  
NNTTAAAAGG G 491

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 407 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..407

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

```
ATTTGTTTCGT GAGGTAACT TTGACTTTAG TCAACGGTAC GAAGCCTGGT GAGAATACGG 60
TTTTAGCGAC TTTCATTGGG ATGTATTTAG GCCGGTCGGA TAAGCTGTTG ACCGTNATGA 120
ACCGGGATTT CCCGGAGTTG AAGCTGAAGA AAACCGATTN TACCGAGATG AGATGGATCG 180
ATTCGGTTCT GTTTTGGGAC GATTATCCGG TTGGTACACC GACTTCTGTG CTACTAAATC 240
CGCTAGTCGC AAAAAAGTTG TTCATGAAAC GAAAATCGGA CTACGTGAAG CGTCTNATTT 300
TCGAGAACCC GATCTCNGT TTGATACTCA AGAAATTTGT AGAGGTTNNG AAAGTTAAAA 360
TNAATTTGGA TCCGCATTNN GGNANNNATG GTGAAACCCC NNGTTNT 407
```

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 360 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..360

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

ACGGCGTCGT ATTGGCCTAC AAAATAAACC TTGTTGAAGT CCCAGAAAAC GTCACCGTTT 60  
TCAGAATCTC CCGGACGTTA GAACAAAATG CGACGGATAT CATTACCGG TGGCAACAAG 120  
TTGCACCGAA GCTTCCCGAC GAGCTTTTCA TAAGANCAGT CATTGACGTA NAAACGGCAC 180  
TGTTTCATNN CTCAAAGAC CGTCAGACAA CATTATAGC AATGTTTCTA GGAGACACGN 240  
CAACTCTACT GTCGATATTA AACCGGAGAT TCCCAGAATT GGGTTTGGTC CGGTCTGACT 300  
GTACCGNAAC AAGCNNTTGG ATCCAATCTG TGCTATTTTT GGGACAAATA TCCCAGGTTG 360

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 427 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..427

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

TCTTCACTGT CACCAAAACG TTAGAACAAG ACGCAAGATT GAAGACTATT TCTAAGTGGC 60  
AACAAATTTC ATCCAAGATT ATTGAAGAGA TACACATCCG AGTGGTACTC AGAGCAGCTG 120  
GAAATGATGG AAACAAGACT GTGACAATGA CCTACCTAGG TCAGTTTCTT GGCGAGAAAG 180  
GCACCTTGCT GAAGGTTATG GAGAAGGCTT TTCCAGAACT AGGGTTAACT CAAAAGGATT 240  
GTACTGAAAT GAGCTGGATT GAAGCCGCCC TTTTCCATGG TGGRTTTCCT ACAGGKTCTC 300  
CTATTGAAAT TTTGCTTMAG CTCAAGTCGC CTYTAGGAAA AGRTTWCTTC AAAGCAACGK 360  
CGGATTTCGT TAAAGAACCT WTTCTGTGA TAGGGCTCAA AGGAATATTC AAAAGATTGA 420  
TTGAAGG 427

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 437 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana

(B) STRAIN: ecotype Columbia

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..437

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

```
GTTGTACTAT CATNGAAGAT TAAGTTAGTC GATGTTCCGT CCACGGTCAC CGNGTTTAA 60
GTCCAGAAAC ATNAGGAGAA AGAGGCCGTT AGGNTCATCA ACAAGTGGCA GTATGTTGCG 120
GATAAGGTCC CTGAAGATCT TTTTCATCAGC GCAACGTTGG NGAGATCAAA CGGAAACTCT 180
GTGCAGGCTT TGTTTACTGG ACTCTATCTT GGNCCGGTGA ATAATNTCTT GGCCTTGATG 240
GAAGAAAAGT TTCCAGANTT AGGTCTTGAT ATCCAAGNCT GCACAGAGAT GAGTTGGGCT 300
GAATCTGCAC TCTGGTNTNC TGNTTTCNCT AAAGGAGAGN CTCCTTGGGT GTTCCNCGCG 360
GATCGGNAGC GGNCAATTTN TGNCTTTCA AGGGGAAAGN CGGCTTTTTN CAAGAACCCG 420
NTACCCGGGG TTCAATT 437
```

(2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 441 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..441

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

```
GCGGACCCTA TAGATCANNA TGTGCTACTG ANAGAAGAGG AAGCCAAGAA CAAGCCGGAG 60
ACAGATAAAT ATCTGAAATG GGNCGATANC GTTTACGAAT TTATGACNCC ATATGTTTCG 120
```



AAATCTCCAA GAGGAGCTTA TGTCAATTC AAGGATATGG ATTTGGGTAT GTATCTTGGA 180  
AAGAAGAAGA CAAAGTACGA GGAAGGAAAG AGTTGGGGAG TGAAGTATTT CAAGAACAAT 240  
TTCGAGAGAT TGGTGAGAGT GAAGACTAGG GTTGATCCAA CAGATTTCTT CTGCGATGAA 300  
CAGAGCATTC CTCTGGTGAA CAAAGTTACC TGAAGATATC ATTTGAAGTT TTTTATTAGT 360  
CCCTTTTCTC TGTGAAATCA TCTGTGCGTG TTGAATATTA TCGTCAAGT GTGTAACCTA 420  
TGTGTGTGAT TGTGAATTGT G 441

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 502 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..502

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

CTGGCTTAAC ACAACGTCGT TTTGGGCCAA TTACCCGGCG GGTACACCCA AGAGCATCCT 60  
TCTAGATAGG CCTCCGACGA ATTCAGTGTC ATTTAAGAGT AAATCGGATT TTGTCAAAAA 120  
ACCAATACCC AAAAAAGGTT TAGAGAAGCT TTGGAAGACA ATGTTTAAAT TCAACAGTAG 180  
CGTCTCGTTG CAATTCAACC CTTACGGTGG AGTGATGGAC CGGATTCCGG CAACGGCCAC 240  
CGCTTTTCCT CATCGGAAAG GAAACTTGTT CAAGGTTCAA TACNCTACGA TGTGGTTTGA 300  
CGCAAACGCC ACACAGAGTA GCCNGGCTAT GATGAATGAG CTTTTTGAGG TGGCGGGACC 360  
GTACGTGNGT CAAGTAAACC CGAGANANGG CTTCTTTTAA NTTAGAGNC CATCGNTNTT 420  
NGGAGCAANN CCAAGTGGGG GGGNCCAACC GGGGGNTNAA ANCNAGNTC TTNGGGGGCC 480  
CAGAATTTCC TTNGGGGAAT TT 502

(2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana  
(B) STRAIN: ecotype Columbia

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 2..400

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

```
NGGGAATTGC NCGAGGNAAG TTGTACCCAA TTCCTGGACC ACCATTGGTT TCCCAAGAAN 60
CCCAGACAA CCGTTTTTCA ATNACCGTGA TGTGATTG GGTATTAATT CTCATAATGG 120
TAAATCAGT AGTTATGTGG AAGGTAAACG TTACGGGAAG AAGTATTCG CAGGTAATTT 180
CGAGAGATTG GTGAAGATTA AGACGAGAGT TGATAGTGGT AATTTCTTTA GGAACGAACA 240
GAGTATTCCT GTGTTACCAT AAGTGATTTT ATTTGATTAT TGGTTAGTGA AATTTGTTGT 300
TGTATAATGA TTATATGTCG TATTTTTATT TATTATTAGT AATTTATAAA GTTTGATATT 360
AAATACAAAT AGTATAATAA GATAGTTTCT TTAGTAAAA 400
```

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 383 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana  
(B) STRAIN: ecotype Columbia

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 2..383

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

```
CAACTCTAAT GGGAACACCT ACTTCGATCG AATGTCGATG GGGGAAGAGC TTTTCTGGGC 60
GGTTCGAGGA GGTGGAGCCG CGAGTTTCGG CATCGTGATG GGATACAAA TCCGGTTGGT 120
```

TCCGGTTCCG GAGAAAGTTA CGGTTTTTAG CGTCGGAAAA ACCGTCGGAG AAGGAGCCGT 180  
 TGATCTTATA ATGAAGTGGC AGAACTTCTC TCATAGTACG GNTCGGAATT TTTTTGTGAA 240  
 GCTGANTTTT GANTTTAGTC AACGGTGCAA AGCCGGGTGA AAAAAAGGTT TTAGNGNCTT 300  
 TCANTTTTGGN TGNAANCTTG GGGGTTTTAT NAGAACGGTT AACCGGGATT NANCCCGNGT 360  
 TTTCCCGGGG TTAAACCTT NGG 383

(2) INFORMATION FOR SEQ ID NO: 36:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 354 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Arabidopsis thaliana
  - (B) STRAIN: ecotype Columbia
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 1..354
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

ATCAATGTTC TTACTAAACG TACACGAGCA TCGTTGGCTT TCAAGGCTAA ATCTGATTTT 60  
 NTTCAAGAAC CGATNCCTAA AACCGCGATT TCGAAGCTTT GGAGACGGTT GCAAGAACCG 120  
 GAAGCAGAGC ATGCTCAGCT AATTTCACN CCATTGCTG GTAAATGAG TNAGATTGCA 180  
 GATTACGAAA CACCATTTC GCATAGGAAG GGGAATATAT ATNAGATTCA GTACTTGAAT 240  
 TACTGGAGAG GAGACGTGAA AGAGAAGTAT ATTGAGATNG GTGGAGGAGA GTTTACGGTT 300  
 GNTATNAGTA AGTTTTTTGG CGAAGTNTNC CNAGAGGNGN CTTNNTNTAA ACCT 354

(2) INFORMATION FOR SEQ ID NO: 37:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 403 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..403

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

```
TTTTTTAGTA CACTAATAAT CAAATGGAAT GAGAAATGAA GCCACAAAAG TATCTGCAAT 60
CAAAATATCC TGCTATCTCC ATCTCAAGCT CTCAATAGTA TCCTCTCCGA AAGTGAAATC 120
AACATTTCAA ACTCTATTTC TTGGTGAAT CGATAGACTG ATTCCTCTGA TGAACCAGAA 180
GTTTCCGGAA CTCGGCTTAC GATCTCAAGA CTGTTCGGAA ATGAGCTGGA TCGAATCGAT 240
AATGTTCTTC AACTGGAGAT CAGGACAGCC GTTAGAGATT TTGCTCAACA GAGACCTAAG 300
GATTGAGGA TCAGTATTTC AAAGCAAAGT CAGGATTATG GTTCAAAAAC CCGTTCCTGA 360
AAACGTTTTT CGAAGAGGTA TCCAAGGGGT TTCTCGAGCA AGT 403
```

(2) INFORMATION FOR SEQ ID NO: 38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 260 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..260

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

```
GAGATGAGTT GGATTAANTC TGTACTCTGG TTTGCTGATT TCCCTAAAGG AGAATCTCTT 60
NGTGTTCTCA CGAATCGTAA GCGTACATCT CTATCTTTNA AAGGCAAAGA TGATTTTATC 120
CAAGAACCGA TACCCGAGGC TGCAATTNAA GAGATATGGA GGCGATTAGA AGCCCCCNAG 180
GCTCGGCTTG GAAAGATCAT ATTA ACTCCA TTTGGTGGGA AAATNAGTGA AATGGCAGAG 240
TACGTANCAC CATTCCACA 260
```

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 605 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..605

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

```
CTCTTGCATA TTCGCTGCAA GGATGGGAAA TTCAAAACCA CTCCCTACAA TTTTGTGTAT 60
TATAGTTTCA GTCTTGTATT TTTAATTCTA TTGCATAACA CCAACTTCTT CATCAGCCTC 120
CATCCAAGAT CAATTCATAA ACTGTGTCAA AAGAAACACA CATGTTTCTT TTCCACTCGA 180
GAAAACGTTA TTCACCCCTG CGAAAAACGT CTCTTTGTTC AACCAAGTCC TTGANTCGAC 240
GGCTCAAAAT CTCCAGTTCT TGGCAAAATC CATGCCTAAA CCGGGRTTCA TATTCAGACC 300
GATTCACCAG TCTCAAGTCC AAGSTCCAT CATTTGTTCa AMGRAACTCG GGNTTCATTT 360
TNGTGTTTGA NGTGGCGGTC ACGATTTTCG AGGCCTTTGT NTTTATGTTT CACGGTTTGA 420
AAAAACCGTT TATATTACTC GGCCTGTCAA ANTTGNANNC AAAATCANAT GTTGGATATT 480
GNATTCCAAA TAGGTNCTTG GGGTNAACCT GGTGGCTANC GTTTGGTGAG CTTTACTTTT 540
CAAGAATTTG CANGNGGANG TGCAAAGATT CCATGGGATT TCCCGGGGGG TTTNTTGCAC 600
AATGT 605
```

(2) INFORMATION FOR SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 464 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*

(B) STRAIN: ecotype Columbia

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 2..464

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

```
AACACAAAAC TCTTCCATTT GGCTTCTCTC TTGCATATTC GTTGCAAGGA TGGGAAATTC 60
AAAACCACTC CCTACAATTN CTTGTATTAT CGTTTCAGTC TTGTATTTTN NATTCTATTG 120
CATAACACCA ACTTCTTCAT CAGCCTCCAT CCAAGNTCAA TTCATAAACT GTGTCAAAAG 180
GAACACACAT GTTTCCTTTT CACTCGAGNA AACGGTATTC ACTCCTGCGG AAAACGGCTC 240
TNTTATTCAA CGGGTCCNTG AATCGACGGG TCAAAATCTC CAGTTCTTGG NAAAATCCAT 300
GNCTAAACCG GGGTTCATAT TCAGGCCGGT TCACCACTCT CAAGTCCAAG NTTCCATCAT 360
TTGTTCAAAG GAACTCGGGA TTCATTTCCG CGNTAGAAGT GGCGGGCANN GGTTTCGGGG 420
CCTGTCTNTT GNTTANGGGN AGGAAAACCG GTTNTATTNC TCGG 464
```

(2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 386 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Arabidopsis thaliana*

(B) STRAIN: ecotype Columbia

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..386

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

```
TCGGGAGCCC ANGNTAAATT ANNTGAAAAT GGGGNCGNAT ANCCGTTTAC NGAATTTTAT 60
GACNCCCAAT ATGTTTCGAA ATCTCAAAGA NNGGGANCTT ATGTCAATTT CAAGGATATG 120
GATTTGGGTA TGTATCTTGG AAAGNAGAAG ACAAAGTACG AGGAAGGAAA GAGTTGGGGA 180
GTGAAGTATT TCAAGAACAA TTTCGAGAGA TTGGTGAGAG TGAAGACTAG GGTTGATCCN 240
ACAGATTTTCN TCTGCGATGA ACAGAGCATT CCTCTGGTGN ACAAAGTTAC CTGAAGATAT 300
CATTTGAAGT TTTTATTAG TCCCTTTTCT CTGTGAAATC ATCTGTGCGT GTTGAATANT 360
```

## (2) INFORMATION FOR SEQ ID NO: 42:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 377 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..377

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

TACCATAGGG AGGTGGTGNA AGATTTTGTA TGTAGNCTTA GGGGAAGGCG AGTAGTATGG 60  
TGGTGGTGGG GAGCTGTAAA CGTATGGTGG TGGTGGAGAT TTGTATGTGG GCTGGTTAAC 120  
TTCATTGAAG CTAAATCTG GGGACCTAAG TACTTCAAAG GCAATTTTGA CAGATTGGTG 180  
AAGATTAAAA CCAAGGTTGA TCCAGAGAAC TTCTTCAGGC ACGAGCAGAG TATCCACCT 240  
ATGCCCTACT AGAAGCTAGG TTCATGAAAC CAATAACATT ATCAAAAATA AGAATAAATG 300  
ATAATTGTAT ACAACATGAT TCGTCTTTCT TTATTTTCTAGA CAATGTGGAC ACTACTCTAA 360  
ATAAAATGTC ATTTACC 377

## (2) INFORMATION FOR SEQ ID NO: 43:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 377 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

## (ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION: 1..377

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

TACCATAGGG AGGTGGTGNA AGATTTTGTG TGTAGNCTTA GGGGAAGGCG AGTAGTATGG 60  
TGGTGGTGGG GAGCTGTAAA CGTATGGTGG TGGTGGAGAT TTGTATGTGG GCTGGTTAAC 120  
TTCATTGAAG CTAAAATCTG GGGACCTAAG TACTTCAAAG GCAATTTTGA CAGATTGGTG 180  
AAGATTAAAA CCAAGGTTGA TCCAGAGAAC TTCTTCAGGC ACGAGCAGAG TATCCACCT 240  
ATGCCCTACT AGAAGCTAGG TTCATGAAAC CAATAACATT ATCAAAAATA AGAATAAATG 300  
ATAATTGTAT ACAACATGAT TCGTCTTTCT TTATTTTCAGA CAATGTGGAC ACTACTCTAA 360  
ATAAAATGTC ATTTACC 377

(2) INFORMATION FOR SEQ ID NO: 44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 346 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*  
(B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION: 2..346

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

GAGCTGTGGA TATGGTCACA AATGGCAATC GGTTGGTCCG AAAACTGATC CGAATCTTTT 60  
TATGAGAATN TTGATTCAAC CAGTGACGAG GAAGAAGGTA AAGACTGTGA GAGCTTCTNT 120  
GGTTGCCCTN TTTTNAGGCN AGACAGATGA AGTTTTTGCT TTCCTTAGTA AGGAGTTTCC 180  
TGAATTGGGT TTAAAGAAGG AGAATTNTTC GGAGATGACT TGGTTTCANT CTGCTTTATG 240  
GTGGGACAAT CGTCTTAATG CTAATCAGGT TGATCCTAAA GTNTTCTTG ATCGGAATCT 300  
CGATACCTCG AGTTTCGGTA AGAGGAAATC GGATTACGTC GCGACT 346

(2) INFORMATION FOR SEQ ID NO: 45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 261 base pairs



- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..261

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

```
ATGGGGTGAG ACTTATTTCA AAGGTAATTT CAAGAGATTA GGTTTGTTA AAGGGAAGNT 60
TGATCCAACA AATTTCTTCA GGAACGAACA GAGTATTCCT CCTCTGTTTT GAGTCCTCAA 120
TACAAAACCA GATATAAAAG ATGTCATTTT ATTTTTTCAA TTATAATAGA TAATGTAAGT 180
TTCTGCTACA ATTGTAAAAG TGAGATGTAC CCAATACGGT TTAAGCGGAC CGAGAATAGT 240
CAATTCAAAG ACCAAATTCT G 261
```

(2) INFORMATION FOR SEQ ID NO: 46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 478 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..478

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

```
GCTCAAAGGA CTAACCATGA AAACCTCCTC AAGTGTCTCT CTCACCGANT CAACGAGGAC 60
GACTCAAGAN TTATACACAC ATCAAAGAT CCTTCGTATT TNTCAATCTT GATTCTTCC 120
ATACAAAATC CAAGTTTCTC TGTTCTCGAA ACACCTAAAC CGGTTTCAAT CATCACTCCG 180
```

GTTCAAGCCA CCGATGTTCA ATCTACGNTT AAATNCGCAC GGNCTTCACG GGTATACACA 240  
 ATCAGGGCTA GGAGTGGTNG TCATGACTAC GGAGGTTTAT CTTTACATTG GCTTAAAAAN 300  
 CANNCCGTTT GTTNNTCATT GATTTNNAGA AATCTTCCGG GCTTATTTAA CATNTAAGAT 360  
 GTTTGATAAN CCGGNNCCNG TTTGGGGTTC AAATCCCGGT GGCTTACAAA NTTNGGGGGA 420  
 ATTGTNCCTA TGAGGTTTGG AAAATTAANG CAAAATNTTT TGGGCCTTCC CGGCCGGT 478

(2) INFORMATION FOR SEQ ID NO: 47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 579 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..579

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

GGCCGTTAGG ATCATCAAGA AATGGCAATA TGCTGCAGAT AAGGTTCTTG ATGATCTTTT 60  
 CATTAGGACA ACATTGGAGA GATCAAACAA GAACGCAGTA CACGCTTTGT TCACTGGACT 120  
 ATATATTGGT CCGGTGAACA ATCTATTGGC GTTGATGGAA GAAAAGTTTC CGGAAGTAGG 180  
 TCTTGAGAAA GAAGGTTGTG AAGAGATGAG TTGGATTGAG TCTGTACTCT GGTTCGCTGA 240  
 TTTCCCTAAA GGAGAATCTC TTGGTGTCT CACGAATCGT GAGCGTACAT CTCTATCTTT 300  
 CAAAGGCAAA GATGATTTTG TCCAAGAACC GATACCCGAG GCTGCAATTC AAGAGATATG 360  
 GAGGCGATTA GAAGCCCCCG AGGCTCGGCT TGGAAAGATC ATATTAATCT CATTGGGGTG 420  
 NGGNAAAATG AGTGAAATGG CAGAGNCCGA ACCACCAATT CCCACANNCG AGGGAGGGGGA 480  
 ACCCCTNTGN GGNTCAGAAT GTGGTTCCTG GNNNNNAAGN GGGNGCCAGN ACCAANCCGG 540  
 GNCNGTAAAN CNTGNAATGG GCCNAACCCG TNCCGGATT 579

(2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 252 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Oryza sativa*  
(B) STRAIN: Nipponbare, subsp. japonica  
(D) DEVELOPMENTAL STAGE: etiolated shoot (8 days old)

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 3..252

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

```
TGTCCTGGAA GGTCCGCCTC GTGCAGGTTN CGACGACGGT GACGGTGTTC GTCGTCGGGA 60
GGAACGTCGA CCAGGGCGCC GCNGACGTCG TCGCCAGATG GCAAGACGTC GCGCCGAGCC 120
TCCCTCCCGA GCTCACCATA CGGGTGATCG TNCGAGGGCA GCGCGCCACG TTCCAGTCGC 180
TGTACCTCGG CTCGTGCGCC GACCTGGTGC CGACGATGAG CAGCATGTTC CCGGAGCTCG 240
GGATGACGAT TG 252
```

(2) INFORMATION FOR SEQ ID NO: 49:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactuca sativa*  
(B) STRAIN: lollo bionda

(ix) FEATURE:

(A) NAME/KEY: Modified-site  
(B) LOCATION: 12  
(D) OTHER INFORMATION: /label= Ambiguous  
/note= "Xaa = Cys or Ser"

(ix) FEATURE:

(A) NAME/KEY: Modified-site  
(B) LOCATION: 20..21  
(D) OTHER INFORMATION: /label= ambiguous  
/note= "Xaa-Xaa probably is Ser-Phe"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

Thr Ser Thr Ser Ile Ile Asp Arg Phe Thr Gln Xaa Leu Asn Asn Arg  
1                      5                      10                      15  
  
Ala Asp Pro Xaa Xaa  
                    20

(2) INFORMATION FOR SEQ ID NO: 50:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Lactuca sativa
- (B) STRAIN: lollo bionda

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 1
- (D) OTHER INFORMATION: /label= ambiguous  
/note= "Xaa = probably Ser"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 3
- (D) OTHER INFORMATION: /label= unknown

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 5
- (D) OTHER INFORMATION: /label= ambiguous  
/note= "Xaa = probably Ser"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 12
- (D) OTHER INFORMATION: /label= ambiguous  
/note= "Xaa = probably Trp"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 24
- (D) OTHER INFORMATION: /label= ambiguous  
/note= "Xaa = probably Tyr"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

Xaa Ile Xaa Val Xaa Ile Glu Asp Glu Thr Ala Xaa Val Gln Ala Gly  
1                      5                      10                      15  
  
Ala Thr Leu Gly Glu Val Tyr Xaa  
                    20

(2) INFORMATION FOR SEQ ID NO: 51:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 14 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Lactuca sativa
  - (B) STRAIN: lollo bionda
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:  

Ala	Asp	Pro	Ser	Phe	Pro	Leu	Ser	Gly	Gln	Leu	Tyr	Tyr	Pro
1				5					10				

(2) INFORMATION FOR SEQ ID NO: 52:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 32 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

ACTTCTACTT CTATTATTGA TAGGTTTACT CA

32

(2) INFORMATION FOR SEQ ID NO: 53:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 405 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Lactuca sativa
  - (B) STRAIN: lollo bionda
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 1..405

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

ACT TCT ACT TCT ATT ATT GAT AGG TTT ACT CAA TGT CTA AAC AAC CGA	48
Thr Ser Thr Ser Ile Ile Asp Arg Phe Thr Gln Cys Leu Asn Asn Arg	
1 5 10 15	
GCT GAC CCT TCT TTC CCG CTC AGT GGA CAA CTT TAC ACT CCC GAT AAC	96
Ala Asp Pro Ser Phe Pro Leu Ser Gly Gln Leu Tyr Thr Pro Asp Asn	
20 25 30	
TCC TCT TTT CCA TCC GTC TTG CAA GCT TAC ATC CGG AAC CTC CGA TTC	144
Ser Ser Phe Pro Ser Val Leu Gln Ala Tyr Ile Arg Asn Leu Arg Phe	
35 40 45	
AAT GAA TCC ACG ACT CCC AAA CCC ATC TTA ATC ATC ACC GCC TTA CAC	192
Asn Glu Ser Thr Thr Pro Lys Pro Ile Leu Ile Ile Thr Ala Leu His	
50 55 60	
CCT TCA CAC ATT CAA GCA GCT GTT GTG TGC GCC AAA ACA CAC CGC CTG	240
Pro Ser His Ile Gln Ala Ala Val Val Cys Ala Lys Thr His Arg Leu	
65 70 75 80	
CTA ATG AAA ACC AGA AGC GGA GGC CAT GAT TAT GAG GGG CTT TCC TAT	288
Leu Met Lys Thr Arg Ser Gly Gly His Asp Tyr Glu Gly Leu Ser Tyr	
85 90 95	
GTG ACC AAT TCG AAC CAA CCC TTT TTT GTT GTT GAC ATG TTC AAC TTA	336
Val Thr Asn Ser Asn Gln Pro Phe Phe Val Val Asp Met Phe Asn Leu	
100 105 110	
CGC TCC ATA AAC GTG AGT ATT GAA GAT GAA ACT GCA TGG GTC CAA GCC	384
Arg Ser Ile Asn Val Ser Ile Glu Asp Glu Thr Ala Trp Val Gln Ala	
115 120 125	
GGC GCC ACC CTC GGA GAA GTT	405
Gly Ala Thr Leu Gly Glu Val	
130 135	

(2) INFORMATION FOR SEQ ID NO: 54:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 135 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

Thr Ser Thr Ser Ile Ile Asp Arg Phe Thr Gln Cys Leu Asn Asn Arg	
1 5 10 15	
Ala Asp Pro Ser Phe Pro Leu Ser Gly Gln Leu Tyr Thr Pro Asp Asn	
20 25 30	
Ser Ser Phe Pro Ser Val Leu Gln Ala Tyr Ile Arg Asn Leu Arg Phe	
35 40 45	
Asn Glu Ser Thr Thr Pro Lys Pro Ile Leu Ile Ile Thr Ala Leu His	
50 55 60	

Pro	Ser	His	Ile	Gln	Ala	Ala	Val	Val	Cys	Ala	Lys	Thr	His	Arg	Leu
65					70					75					80
Leu	Met	Lys	Thr	Arg	Ser	Gly	Gly	His	Asp	Tyr	Glu	Gly	Leu	Ser	Tyr
				85					90					95	
Val	Thr	Asn	Ser	Asn	Gln	Pro	Phe	Phe	Val	Val	Asp	Met	Phe	Asn	Leu
			100					105						110	
Arg	Ser	Ile	Asn	Val	Ser	Ile	Glu	Asp	Glu	Thr	Ala	Trp	Val	Gln	Ala
		115					120					125			
Gly	Ala	Thr	Leu	Gly	Glu	Val									
	130					135									

(2) INFORMATION FOR SEQ ID NO: 55:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

CACGTTTATG GAGCGTAAGT TGAAC

25

(2) INFORMATION FOR SEQ ID NO: 56:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

CACCCTTCAC ACATTCAAGC AGC

23

(2) INFORMATION FOR SEQ ID NO: 57:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1981 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Lactuca sativa
- (B) STRAIN: lollo bionda

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 7..1626

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 372
- (D) OTHER INFORMATION: location 372 may be "C" or "G"

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 379
- (D) OTHER INFORMATION: location 379 may be "A" or "G"

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 786
- (D) OTHER INFORMATION: location 786 may be "C" or "T"

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 1105...1106
- (D) OTHER INFORMATION: location 1105...1106 may be "AG", "GA", "GG" or "AA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

ACAAAA ATG GCA ATT ACC TAT TCT TTC AAC TTC AAA TCT TAT ATT TTT	48
Met Ala Ile Thr Tyr Ser Phe Asn Phe Lys Ser Tyr Ile Phe	
1 5 10	
CCT CTC CTC CTT GTC TTG CTC TCT ACC CAT TCA TCA GCG ACT TCA ACT	96
Pro Leu Leu Leu Val Leu Leu Ser Thr His Ser Ser Ala Thr Ser Thr	
15 20 25 30	
TCC ATT ATA GAT CGC TTC ACC CAA TGT CTA AAC AAC CGA GCT GAC CCT	144
Ser Ile Ile Asp Arg Phe Thr Gln Cys Leu Asn Asn Arg Ala Asp Pro	
35 40 45	
TCT TTC CCG CTC AGT GGA CAA CTT TAC ACT CCC GAT AAC TCC TCT TTT	192
Ser Phe Pro Leu Ser Gly Gln Leu Tyr Thr Pro Asp Asn Ser Ser Phe	
50 55 60	
CCA TCC GTC TTG CAA GCT TAC ATC CGG AAC CTC CGA TTC AAT GAA TCC	240
Pro Ser Val Leu Gln Ala Tyr Ile Arg Asn Leu Arg Phe Asn Glu Ser	
65 70 75	
ACG ACT CCC AAA CCC ATC TTA ATC ATC ACC GCC TTA CAC CCT TCA CAC	288
Thr Thr Pro Lys Pro Ile Leu Ile Ile Thr Ala Leu His Pro Ser His	
80 85 90	
ATT CAA GCA GCT GTT GTG TGC GCC AAA ACA CAC CGC CTG CTA ATG AAA	336
Ile Gln Ala Ala Val Val Cys Ala Lys Thr His Arg Leu Leu Met Lys	
95 100 105 110	



ACC AGA AGC GGA GGC CAT GAT TAT GAG GGG CTT TCS TAT GTG RCC AAT	384
Thr Arg Ser Gly Gly His Asp Tyr Glu Gly Leu Ser Tyr Val Xaa Asn	
115 120 125	
TCG AAC CAA CCC TTT TTT GTT GTT GAC ATG TTC AAC TTA CGC TCC ATA	432
Ser Asn Gln Pro Phe Phe Val Val Asp Met Phe Asn Leu Arg Ser Ile	
130 135 140	
AAC GTG AGT ATT GAA GAT GAA ACT GCA TGG GTC CAA GCT GGT GCG ACT	480
Asn Val Ser Ile Glu Asp Glu Thr Ala Trp Val Gln Ala Gly Ala Thr	
145 150 155	
CTT GGT GAA GTC TAC TAC CGA ATA GCA GAG AAA AGC AAC AGT CAT GCT	528
Leu Gly Glu Val Tyr Tyr Arg Ile Ala Glu Lys Ser Asn Ser His Ala	
160 165 170	
TTT CCG GCT GGC GTT TGC CCT ACT GTT GGA GTT GGT GGC CAT TTT AGT	576
Phe Pro Ala Gly Val Cys Pro Thr Val Gly Val Gly Gly His Phe Ser	
175 180 185 190	
GGT GGT GGT TAT GGT AAC TTG ATG GGA AAA TAC GGC CTT TCT GTT GAC	624
Gly Gly Gly Tyr Gly Asn Leu Met Gly Lys Tyr Gly Leu Ser Val Asp	
195 200 205	
AAT ATT GTC GAT GCT CAG TTA ATC GAT GTG AAT GGT AAA CTT CTG AAT	672
Asn Ile Val Asp Ala Gln Leu Ile Asp Val Asn Gly Lys Leu Leu Asn	
210 215 220	
CGG AAA TCA ATG GGT GAA GAT CTT TTT TGG GCC ATC ACA GGT GGT GGT	720
Arg Lys Ser Met Gly Glu Asp Leu Phe Trp Ala Ile Thr Gly Gly Gly	
225 230 235	
GGT GTC AGC TTT GGT GTG GTT GTA GCG TAC AAG ATC AAA CTG GTT CGT	768
Gly Val Ser Phe Gly Val Val Val Ala Tyr Lys Ile Lys Leu Val Arg	
240 245 250	
GTT CCT ACC ACT GTG ACY GTT TTT AAC GTA CAA AGA ACA TCC GAG CAG	816
Val Pro Thr Thr Val Thr Val Phe Asn Val Gln Arg Thr Ser Glu Gln	
255 260 265 270	
AAC CTA AGC ACC ATA GCC CAC CGA TGG ATA CAA GTT GCG GAT AAG CTC	864
Asn Leu Ser Thr Ile Ala His Arg Trp Ile Gln Val Ala Asp Lys Leu	
275 280 285	
GAT AAT GAC CTT TTC CTT CGA ATG ACC TTT AAC GTG ATA AAC AAC ACA	912
Asp Asn Asp Leu Phe Leu Arg Met Thr Phe Asn Val Ile Asn Asn Thr	
290 295 300	
AAT GGC GAA AAG ACG ATA CGT GGT TTG TTT CCA ACA CTG TAC CTC GGA	960
Asn Gly Glu Lys Thr Ile Arg Gly Leu Phe Pro Thr Leu Tyr Leu Gly	
305 310 315	
AAC TCT ACC GCT CTT GTT GCC CTC CTG AAC AAG GAT TTC CCT GAA TTA	1008
Asn Ser Thr Ala Leu Val Ala Leu Leu Asn Lys Asp Phe Pro Glu Leu	
320 325 330	
GGT GTA GAA ATT TCA GAT TGT ATT GAA ATG AGT TGG ATC GAG TCT GTT	1056
Gly Val Glu Ile Ser Asp Cys Ile Glu Met Ser Trp Ile Glu Ser Val	
335 340 345 350	

CTT TTC TAC ACA AAC TTC CCC ATT GGT ACT CCG ACC ACT GCT CTT CTA	1104
Leu Phe Tyr Thr Asn Phe Pro Ile Gly Thr Pro Thr Thr Ala Leu Leu	
355 360 365	
RRC CGT ACA CCT CAA AGA CTA AAC CCA TTC AAA ATC AAA TCT GAT TAC	1152
Xaa Arg Thr Pro Gln Arg Leu Asn Pro Phe Lys Ile Lys Ser Asp Tyr	
370 375 380	
GTA AAA AAC ACT ATT TCC AAA CAG GGA TTC GAA TCC ATA TTT GAA AGG	1200
Val Lys Asn Thr Ile Ser Lys Gln Gly Phe Glu Ser Ile Phe Glu Arg	
385 390 395	
ATG AAA GAA CTC GAA AAC CAA ATG CTA GCT TTC AAC CCT TAT GGT GGA	1248
Met Lys Glu Leu Glu Asn Gln Met Leu Ala Phe Asn Pro Tyr Gly Gly	
400 405 410	
AGA ATG AGC GAA ATT TCC GAA TTT GCA AAG CCT TTT CCC CAT CGA TCA	1296
Arg Met Ser Glu Ile Ser Glu Phe Ala Lys Pro Phe Pro His Arg Ser	
415 420 425 430	
GGG AAT ATA GCG AAG ATC CAA TAC GAA GTA AAC TGG GAT GAA CTT GGC	1344
Gly Asn Ile Ala Lys Ile Gln Tyr Glu Val Asn Trp Asp Glu Leu Gly	
435 440 445	
GTT GAA GCA GCC AAT CGG TAC TTG AAC TTC ACA AGG GTG ATG TAT GAT	1392
Val Glu Ala Ala Asn Arg Tyr Leu Asn Phe Thr Arg Val Met Tyr Asp	
450 455 460	
TAT ATG ACT CCG TTT GTT TCT AAG AAC CCC AGG GAA GCA TTT CTG AAC	1440
Tyr Met Thr Pro Phe Val Ser Lys Asn Pro Arg Glu Ala Phe Leu Asn	
465 470 475	
TAC AGG GAT TTA GAT ATT GGT GTC AAC AGT CAT GGC AAG AAT GCT TAC	1488
Tyr Arg Asp Leu Asp Ile Gly Val Asn Ser His Gly Lys Asn Ala Tyr	
480 485 490	
GGT GAA GGA ATG GTT TAT GGG CAC AAG TAT TTC AAA GAG ACG AAT TAT	1536
Gly Glu Gly Met Val Tyr Gly His Lys Tyr Phe Lys Glu Thr Asn Tyr	
495 500 505 510	
AAG AGG CTA ACG ATG GTG AAG ACG AGG GTT GAT CCT AGC AAT TTT TTT	1584
Lys Arg Leu Thr Met Val Lys Thr Arg Val Asp Pro Ser Asn Phe Phe	
515 520 525	
AGG AAT GAG CAA AGT ATC CCA ACT TTG TCA TCT TCA TGG AAG	1626
Arg Asn Glu Gln Ser Ile Pro Thr Leu Ser Ser Ser Trp Lys	
530 535 540	
TAAATTCTAA ATTCACTTGT GAAATTGAAT AAAAGTATGG CTTTTTCAAG GTCATGGTAT	1686
CCAGATTCAG ATGATATTGA TATAATTTTG ACTTGATTTT ATACAAACAA AATTATATTA	1746
TATTTTTCTG AATTAGATT TTCCATTCTT TGGAAAAATA TACGAACATT GATGTTGATA	1806
TTTTTAAGAA TTATAGATTT TGAACATTGT GAACAATGAA TAAACCGAGG ACTTCCCTTG	1866
GGTTTTTTTTT ATAAGTATGT AATAGCATGT CTTTAATCAA GATAACCGAT CATTGGATGC	1926
AATTTATTAT TATAAACCTT ATTTAAAAA AAAAAAAAAA AAAAAAAAAA AAAAA	1981

(2) INFORMATION FOR SEQ ID NO: 58:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 540 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

Met	Ala	Ile	Thr	Tyr	Ser	Phe	Asn	Phe	Lys	Ser	Tyr	Ile	Phe	Pro	Leu	
1				5					10					15		
Leu	Leu	Val	Leu	Leu	Ser	Thr	His	Ser	Ser	Ala	Thr	Ser	Thr	Ser	Ile	
			20					25					30			
Ile	Asp	Arg	Phe	Thr	Gln	Cys	Leu	Asn	Asn	Arg	Ala	Asp	Pro	Ser	Phe	
	35						40					45				
Pro	Leu	Ser	Gly	Gln	Leu	Tyr	Thr	Pro	Asp	Asn	Ser	Ser	Phe	Pro	Ser	
	50					55					60					
Val	Leu	Gln	Ala	Tyr	Ile	Arg	Asn	Leu	Arg	Phe	Asn	Glu	Ser	Thr	Thr	
65					70					75					80	
Pro	Lys	Pro	Ile	Leu	Ile	Ile	Thr	Ala	Leu	His	Pro	Ser	His	Ile	Gln	
			85						90					95		
Ala	Ala	Val	Val	Cys	Ala	Lys	Thr	His	Arg	Leu	Leu	Met	Lys	Thr	Arg	
		100						105					110			
Ser	Gly	Gly	His	Asp	Tyr	Glu	Gly	Leu	Ser	Tyr	Val	Thr	Asn	Ser	Asn	
	115						120					125				
Gln	Pro	Phe	Phe	Val	Val	Asp	Met	Phe	Asn	Leu	Arg	Ser	Ile	Asn	Val	
	130					135					140					
Ser	Ile	Glu	Asp	Glu	Thr	Ala	Trp	Val	Gln	Ala	Gly	Ala	Thr	Leu	Gly	
145					150				155					160		
Glu	Val	Tyr	Tyr	Arg	Ile	Ala	Glu	Lys	Ser	Asn	Ser	His	Ala	Phe	Pro	
			165					170					175			
Ala	Gly	Val	Cys	Pro	Thr	Val	Gly	Val	Gly	Gly	His	Phe	Ser	Gly	Gly	
		180					185					190				
Gly	Tyr	Gly	Asn	Leu	Met	Gly	Lys	Tyr	Gly	Leu	Ser	Val	Asp	Asn	Ile	
	195						200					205				
Val	Asp	Ala	Gln	Leu	Ile	Asp	Val	Asn	Gly	Lys	Leu	Leu	Asn	Arg	Lys	
	210					215					220					
Ser	Met	Gly	Glu	Asp	Leu	Phe	Trp	Ala	Ile	Thr	Gly	Gly	Gly	Gly	Val	
225					230				235					240		
Ser	Phe	Gly	Val	Val	Val	Ala	Tyr	Lys	Ile	Lys	Leu	Val	Arg	Val	Pro	
			245					250					255			
Thr	Thr	Val	Thr	Val	Phe	Asn	Val	Gln	Arg	Thr	Ser	Glu	Gln	Asn	Leu	
		260					265						270			

Ser Thr Ile Ala His Arg Trp Ile Gln Val Ala Asp Lys Leu Asp Asn  
 275 280 285  
 Asp Leu Phe Leu Arg Met Thr Phe Asn Val Ile Asn Asn Thr Asn Gly  
 290 295 300  
 Glu Lys Thr Ile Arg Gly Leu Phe Pro Thr Leu Tyr Leu Gly Asn Ser  
 305 310 315 320  
 Thr Ala Leu Val Ala Leu Leu Asn Lys Asp Phe Pro Glu Leu Gly Val  
 325 330 335  
 Glu Ile Ser Asp Cys Ile Glu Met Ser Trp Ile Glu Ser Val Leu Phe  
 340 345 350  
 Tyr Thr Asn Phe Pro Ile Gly Thr Pro Thr Thr Ala Leu Leu Ser Arg  
 355 360 365  
 Thr Pro Gln Arg Leu Asn Pro Phe Lys Ile Lys Ser Asp Tyr Val Lys  
 370 375 380  
 Asn Thr Ile Ser Lys Gln Gly Phe Glu Ser Ile Phe Glu Arg Met Lys  
 385 390 395 400  
 Glu Leu Glu Asn Gln Met Leu Ala Phe Asn Pro Tyr Gly Gly Arg Met  
 405 410 415  
 Ser Glu Ile Ser Glu Phe Ala Lys Pro Phe Pro His Arg Ser Gly Asn  
 420 425 430  
 Ile Ala Lys Ile Gln Tyr Glu Val Asn Trp Asp Glu Leu Gly Val Glu  
 435 440 445  
 Ala Ala Asn Arg Tyr Leu Asn Phe Thr Arg Val Met Tyr Asp Tyr Met  
 450 455 460  
 Thr Pro Phe Val Ser Lys Asn Pro Arg Glu Ala Phe Leu Asn Tyr Arg  
 465 470 475 480  
 Asp Leu Asp Ile Gly Val Asn Ser His Gly Lys Asn Ala Tyr Gly Glu  
 485 490 495  
 Gly Met Val Tyr Gly His Lys Tyr Phe Lys Glu Thr Asn Tyr Lys Arg  
 500 505 510  
 Leu Thr Met Val Lys Thr Arg Val Asp Pro Ser Asn Phe Phe Arg Asn  
 515 520 525  
 Glu Gln Ser Ile Pro Thr Leu Ser Ser Ser Trp Lys  
 530 535 540

(2) INFORMATION FOR SEQ ID NO: 59:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

GGTAATGATC TCCTTTCTTG TTTGACC

27

(2) INFORMATION FOR SEQ ID NO: 60:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 41 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

AGAGCGGCCG CTATATTACA ACTTCTCCAC CATCACTCCT C

41

(2) INFORMATION FOR SEQ ID NO: 61:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

GGTGATGTTA ATGATAATCT CCTC

24

(2) INFORMATION FOR SEQ ID NO: 62:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 40 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

AGAGCGGCCG CTACAATTCC TTCAACATGT AAATTTTCCTC

40

(2) INFORMATION FOR SEQ ID NO: 63:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 36 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

ACTTCCCGTA GAAACTCGGA GACTTTCACA CAATGC

36

(2) INFORMATION FOR SEQ ID NO: 64:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 30 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:

TCCATCCAAG ATCAATTCAT AAAGTGTGTC

30

(2) INFORMATION FOR SEQ ID NO: 65:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 36 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 65:

AGAGCGGCCG CTTTCATGAA CCTAGCTTCT AGTAGG

36

(2) INFORMATION FOR SEQ ID NO: 66:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 37 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:

AGAGCGGCCG CGAAATGGCC CCCCTTTTAA AACGGGG

37

(2) INFORMATION FOR SEQ ID NO: 67:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

AGAGCGGCCG CAAATGATAT CTTCAGGTAA CTTTGTTCAC

40

(2) INFORMATION FOR SEQ ID NO: 68:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:

AGAGCGGCCG CATAATCAAA TAAATACACT TATGGTAACA CAG

43

(2) INFORMATION FOR SEQ ID NO: 69:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:

AGAGCGGCCG CTGGTTTGT ATTGAGGACT CAAAACAG

38

(2) INFORMATION FOR SEQ ID NO: 70:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1757 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Arabidopsis thaliana*  
(B) STRAIN: Colombia

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: join(1..570, 801..1754)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:

ACT TCC CGT AGA AAC TCG GAG ACT TTC ACA CAA TGC CTA ACC TCA AAC	48
Thr Ser Arg Arg Asn Ser Glu Thr Phe Thr Gln Cys Leu Thr Ser Asn	
1 5 10 15	
TCC GAC CCC AAA CAT CCC ATC TCC CCC GCT ATC TTC TTC TCC GGA AAT	96
Ser Asp Pro Lys His Pro Ile Ser Pro Ala Ile Phe Phe Ser Gly Asn	
20 25 30	
GGC TCC TAC TCC TCC GTA TTA CAA GCC AAC ATC CGT AAC CTC CGC TTC	144
Gly Ser Tyr Ser Ser Val Leu Gln Ala Asn Ile Arg Asn Leu Arg Phe	
35 40 45	
AAC ACC ACC TCA ACT CCG AAA CCC TTC CTC ATA ATC GCC GCA ACA CAT	192
Asn Thr Thr Ser Thr Pro Lys Pro Phe Leu Ile Ile Ala Ala Thr His	
50 55 60	
GAA TCC CAT GTG CAA GCC GCG ATT ACT TGC GGG AAA CGC CAC AAC CTT	240
Glu Ser His Val Gln Ala Ala Ile Thr Cys Gly Lys Arg His Asn Leu	
65 70 75 80	
CAG ATG AAA ATC AGA AGT GGA GGC CAC GAC TAC GAT GGC TTG TCA TAC	288
Gln Met Lys Ile Arg Ser Gly Gly His Asp Tyr Asp Gly Leu Ser Tyr	
85 90 95	
GTT ACA TAC TCT GGC AAA CCG TTC TTC GTC CTC GAC ATG TTT AAC CTC	336
Val Thr Tyr Ser Gly Lys Pro Phe Phe Val Leu Asp Met Phe Asn Leu	
100 105 110	
CGT TCG GTG GAT GTC GAT GTG GCA AGT AAG ACC GCG TGG GTC CAA ACC	384
Arg Ser Val Asp Val Asp Val Ala Ser Lys Thr Ala Trp Val Gln Thr	
115 120 125	
GGT GCC ATA CTC GGA GAA GTT TAT TAC TAT ATA TGG GAG AAG AGC AAA	432
Gly Ala Ile Leu Gly Glu Val Tyr Tyr Tyr Ile Trp Glu Lys Ser Lys	
130 135 140	
ACC CTA GCT TAT CCC GCC GGA ATT TGT CCC ACG GTT GGT GTC GGT GGC	480
Thr Leu Ala Tyr Pro Ala Gly Ile Cys Pro Thr Val Gly Val Gly Gly	
145 150 155 160	



CAT ATC AGT GGT GGA GGT TAC GGT AAC ATG ATG AGA AAA TAC GGT CTC	528
His Ile Ser Gly Gly Gly Tyr Gly Asn Met Met Arg Lys Tyr Gly Leu	
165 170 175	
ACC GTA GAT AAT ACC ATC GAT GCA AGA ATG GTC GAC GTT AAT	570
Thr Val Asp Asn Thr Ile Asp Ala Arg Met Val Asp Val Asn	
180 185 190	
GGTATAATTG ATATCTCTAT TTTATATACT AATTAAATTT TATAGTGTGG ATCGGATAGT	630
GATTTTGGTC CATCAATTAA AAACCTGGTG AACATAAAAT TAACCAAGCA ATCAATTTAG	690
ACAAGCAACA TAATCATATA TATTTTTCTT ACATTTGTAT GTACCTGAAT ATTTATATTT	750
ATGTTTATAT GTTCTCACTA TATTTTCACT TTTGTATTTG AAAATTTTTA GGA AAA	806
Gly Lys	
ATT TTG GAT AGA AAA TTG ATG GGA GAA GAT CTC TAC TGG GCA ATA AAC	854
Ile Leu Asp Arg Lys Leu Met Gly Glu Asp Leu Tyr Trp Ala Ile Asn	
195 200 205	
GGA GGA GGA GGA GGG AGC TAC GGC GTC GTA TTG GCC TAC AAA ATA AAC	902
Gly Gly Gly Gly Gly Ser Tyr Gly Val Val Leu Ala Tyr Lys Ile Asn	
210 215 220	
CTT GTT GAA GTC CCA GAA AAC GTC ACC GTT TTC AGA ATC TCC CGG ACG	950
Leu Val Glu Val Pro Glu Asn Val Thr Val Phe Arg Ile Ser Arg Thr	
225 230 235 240	
TTA GAA CAA AAT GCG ACG GAT ATC ATT CAC CGG TGG CAA CAA GTT GCA	998
Leu Glu Gln Asn Ala Thr Asp Ile Ile His Arg Trp Gln Gln Val Ala	
245 250 255	
CCG AAG CTT CCC GAC GAG CTT TTC ATA AGA ACA GTC ATT GAC GTA GTA	1046
Pro Lys Leu Pro Asp Glu Leu Phe Ile Arg Thr Val Ile Asp Val Val	
260 265 270	
AAC GGC ACT GTT TCA TCT CAA AAG ACC GTC AGG ACA ACA TTC ATA GCA	1094
Asn Gly Thr Val Ser Ser Gln Lys Thr Val Arg Thr Thr Phe Ile Ala	
275 280 285	
ATG TTT CTA GGA GAC ACG ACA ACT CTA CTG TCG ATA TTA AAC CGG AGA	1142
Met Phe Leu Gly Asp Thr Thr Thr Leu Leu Ser Ile Leu Asn Arg Arg	
290 295 300	
TTC CCA GAA TTG GGT TTG GTC CGG TCT GAC TGT ACC GAA ACA AGC TGG	1190
Phe Pro Glu Leu Gly Leu Val Arg Ser Asp Cys Thr Glu Thr Ser Trp	
305 310 315 320	
ATC CAA TCT GTG CTA TTC TGG ACA AAT ATC CAA GTT GGT TCG TCG GAG	1238
Ile Gln Ser Val Leu Phe Trp Thr Asn Ile Gln Val Gly Ser Ser Glu	
325 330 335	
ACA CTT CTA CTC CAA AGG AAT CAA CCC GTG AAC TAC CTC AAG AGG AAA	1286
Thr Leu Leu Leu Gln Arg Asn Gln Pro Val Asn Tyr Leu Lys Arg Lys	
340 345 350	
TCA GAT TAC GTA CGT GAA CCG ATT TCA AGA ACC GGT TTA GAG TCA ATT	1334
Ser Asp Tyr Val Arg Glu Pro Ile Ser Arg Thr Gly Leu Glu Ser Ile	
355 360 365	

TGG AAG AAA ATG ATC GAG CTT GAA ATT CCG ACA ATG GCT TTC AAT CCA 1382  
 Trp Lys Lys Met Ile Glu Leu Glu Ile Pro Thr Met Ala Phe Asn Pro  
 370 375 380

TAC GGT GGT GAG ATG GGG AGG ATA TCA TTA CGG GTG ACT CCG TTC CCA 1430  
 Tyr Gly Gly Glu Met Gly Arg Ile Ser Leu Arg Val Thr Pro Phe Pro  
 385 390 395 400

TAC AGA GCC GGT AAT CTC TGG AAG ATT CAG TAC GGT GCG AAT TGG AGA 1478  
 Tyr Arg Ala Gly Asn Leu Trp Lys Ile Gln Tyr Gly Ala Asn Trp Arg  
 405 410 415

GAT GAG ACT TTA ACC GAC CGG TAC ATG GAA TTG ACG AGG AAG TTG TAC 1526  
 Asp Glu Thr Leu Thr Asp Arg Tyr Met Glu Leu Thr Arg Lys Leu Tyr  
 420 425 430

CAA TTC ATG ACA CCA TTT GTT TCC AAG AAT CCG AGA CAA TCG TTT TTC 1574  
 Gln Phe Met Thr Pro Phe Val Ser Lys Asn Pro Arg Gln Ser Phe Phe  
 435 440 445

AAT AAC CGT GAT GTT GAT TTG GGT ATT AAT TCT CAT AAT GGT AAA ATC 1622  
 Asn Asn Arg Asp Val Asp Leu Gly Ile Asn Ser His Asn Gly Lys Ile  
 450 455 460

AGT AGT TAT GTG GAA GGT AAA CGT TAC GGG AAG AAG TAT TTC GCA GGT 1670  
 Ser Ser Tyr Val Glu Gly Lys Arg Tyr Gly Lys Lys Tyr Phe Ala Gly  
 465 470 475 480

AAT TTC GAG AGA TTG GTG AAG ATT AAG ACG AGA GTT GAT AGT GGT AAT 1718  
 Asn Phe Glu Arg Leu Val Lys Ile Lys Thr Arg Val Asp Ser Gly Asn  
 485 490 495

TTC TTT AGG AAC GAA CAC AGT ATT CCT GTG TTA CCA TAA 1757  
 Phe Phe Arg Asn Glu His Ser Ile Pro Val Leu Pro  
 500 505

(2) INFORMATION FOR SEQ ID NO: 71:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 508 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

Thr Ser Arg Arg Asn Ser Glu Thr Phe Thr Gln Cys Leu Thr Ser Asn  
 1 5 10 15

Ser Asp Pro Lys His Pro Ile Ser Pro Ala Ile Phe Phe Ser Gly Asn  
 20 25 30

Gly Ser Tyr Ser Ser Val Leu Gln Ala Asn Ile Arg Asn Leu Arg Phe  
 35 40 45

Asn Thr Thr Ser Thr Pro Lys Pro Phe Leu Ile Ile Ala Ala Thr His  
 50 55 60

Glu	Ser	His	Val	Gln	Ala	Ala	Ile	Thr	Cys	Gly	Lys	Arg	His	Asn	Leu	
65					70					75					80	
Gln	Met	Lys	Ile	Arg	Ser	Gly	Gly	His	Asp	Tyr	Asp	Gly	Leu	Ser	Tyr	
				85					90					95		
Val	Thr	Tyr	Ser	Gly	Lys	Pro	Phe	Phe	Val	Leu	Asp	Met	Phe	Asn	Leu	
			100					105					110			
Arg	Ser	Val	Asp	Val	Asp	Val	Ala	Ser	Lys	Thr	Ala	Trp	Val	Gln	Thr	
		115					120					125				
Gly	Ala	Ile	Leu	Gly	Glu	Val	Tyr	Tyr	Tyr	Ile	Trp	Glu	Lys	Ser	Lys	
	130					135					140					
Thr	Leu	Ala	Tyr	Pro	Ala	Gly	Ile	Cys	Pro	Thr	Val	Gly	Val	Gly	Gly	
145					150					155					160	
His	Ile	Ser	Gly	Gly	Gly	Tyr	Gly	Asn	Met	Met	Arg	Lys	Tyr	Gly	Leu	
			165					170						175		
Thr	Val	Asp	Asn	Thr	Ile	Asp	Ala	Arg	Met	Val	Asp	Val	Asn	Gly	Lys	
		180						185					190			
Ile	Leu	Asp	Arg	Lys	Leu	Met	Gly	Glu	Asp	Leu	Tyr	Trp	Ala	Ile	Asn	
	195						200					205				
Gly	Gly	Gly	Gly	Gly	Ser	Tyr	Gly	Val	Val	Leu	Ala	Tyr	Lys	Ile	Asn	
	210					215					220					
Leu	Val	Glu	Val	Pro	Glu	Asn	Val	Thr	Val	Phe	Arg	Ile	Ser	Arg	Thr	
225					230					235					240	
Leu	Glu	Gln	Asn	Ala	Thr	Asp	Ile	Ile	His	Arg	Trp	Gln	Gln	Val	Ala	
			245						250					255		
Pro	Lys	Leu	Pro	Asp	Glu	Leu	Phe	Ile	Arg	Thr	Val	Ile	Asp	Val	Val	
		260						265					270			
Asn	Gly	Thr	Val	Ser	Ser	Gln	Lys	Thr	Val	Arg	Thr	Thr	Phe	Ile	Ala	
		275					280					285				
Met	Phe	Leu	Gly	Asp	Thr	Thr	Thr	Leu	Leu	Ser	Ile	Leu	Asn	Arg	Arg	
	290					295					300					
Phe	Pro	Glu	Leu	Gly	Leu	Val	Arg	Ser	Asp	Cys	Thr	Glu	Thr	Ser	Trp	
305					310					315					320	
Ile	Gln	Ser	Val	Leu	Phe	Trp	Thr	Asn	Ile	Gln	Val	Gly	Ser	Ser	Glu	
			325					330						335		
Thr	Leu	Leu	Leu	Gln	Arg	Asn	Gln	Pro	Val	Asn	Tyr	Leu	Lys	Arg	Lys	
			340					345					350			
Ser	Asp	Tyr	Val	Arg	Glu	Pro	Ile	Ser	Arg	Thr	Gly	Leu	Glu	Ser	Ile	
		355					360					365				
Trp	Lys	Lys	Met	Ile	Glu	Leu	Glu	Ile	Pro	Thr	Met	Ala	Phe	Asn	Pro	
	370					375					380					

Tyr Gly Gly Glu Met Gly Arg Ile Ser Leu Arg Val Thr Pro Phe Pro  
 385 390 395 400  
 Tyr Arg Ala Gly Asn Leu Trp Lys Ile Gln Tyr Gly Ala Asn Trp Arg  
 405 410 415  
 Asp Glu Thr Leu Thr Asp Arg Tyr Met Glu Leu Thr Arg Lys Leu Tyr  
 420 425 430  
 Gln Phe Met Thr Pro Phe Val Ser Lys Asn Pro Arg Gln Ser Phe Phe  
 435 440 445  
 Asn Asn Arg Asp Val Asp Leu Gly Ile Asn Ser His Asn Gly Lys Ile  
 450 455 460  
 Ser Ser Tyr Val Glu Gly Lys Arg Tyr Gly Lys Lys Tyr Phe Ala Gly  
 465 470 475 480  
 Asn Phe Glu Arg Leu Val Lys Ile Lys Thr Arg Val Asp Ser Gly Asn  
 485 490 495  
 Phe Phe Arg Asn Glu His Ser Ile Pro Val Leu Pro  
 500 505

(2) INFORMATION FOR SEQ ID NO: 72:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1527 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: Colombia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1524

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

ACT TCC CGT AGA AAC TCG GAG ACT TTC ACA CAA TGC CTA ACC TCA AAC	48
Thr Ser Arg Arg Asn Ser Glu Thr Phe Thr Gln Cys Leu Thr Ser Asn	
1 5 10 15	
TCC GAC CCC AAA CAT CCC ATC TCC CCC GCT ATC TTC TTC TCC GGA AAT	96
Ser Asp Pro Lys His Pro Ile Ser Pro Ala Ile Phe Phe Ser Gly Asn	
20 25 30	
GGC TCC TAC TCC TCC GTA TTA CAA GCC AAC ATC CGT AAC CTC CGC TTC	144
Gly Ser Tyr Ser Ser Val Leu Gln Ala Asn Ile Arg Asn Leu Arg Phe	
35 40 45	

AAC ACC ACC TCA ACT CCG AAA CCC TTC CTC ATA ATC GCC GCA ACA CAT	192
Asn Thr Thr Ser Thr Pro Lys Pro Phe Leu Ile Ile Ala Ala Thr His	
50 55 60	
GAA TCC CAT GTG CAA GCC GCG ATT ACT TGC GGG AAA CGC CAC AAC CTT	240
Glu Ser His Val Gln Ala Ala Ile Thr Cys Gly Lys Arg His Asn Leu	
65 70 75 80	
CAG ATG AAA ATC AGA AGT GGA GGC CAC GAC TAC GAT GGC TTG TCA TAC	288
Gln Met Lys Ile Arg Ser Gly Gly His Asp Tyr Asp Gly Leu Ser Tyr	
85 90 95	
GTT ACA TAC TCT GGC AAA CCG TTC TTC GTC CTC GAC ATG TTT AAC CTC	336
Val Thr Tyr Ser Gly Lys Pro Phe Phe Val Leu Asp Met Phe Asn Leu	
100 105 110	
CGT TCG GTG GAT GTC GAC GTG GCA AGT AAG ACC GCG TGG GTC CAA ACC	384
Arg Ser Val Asp Val Asp Val Ala Ser Lys Thr Ala Trp Val Gln Thr	
115 120 125	
GGT GCC ATA CTC GGA GAA GTT TAT TAC TAT ATA TGG GAG AAG AGC AAA	432
Gly Ala Ile Leu Gly Glu Val Tyr Tyr Tyr Ile Trp Glu Lys Ser Lys	
130 135 140	
ACC CTA GCT TAT CCC GCC GGA ATT TGT CCC ACG GTT GGT GTC GGT GGC	480
Thr Leu Ala Tyr Pro Ala Gly Ile Cys Pro Thr Val Gly Val Gly Gly	
145 150 155 160	
CAT ATC AGT GGT GGA GGT TAC GGT AAC ATG ATG AGA AAA TAC GGT CTC	528
His Ile Ser Gly Gly Gly Tyr Gly Asn Met Met Arg Lys Tyr Gly Leu	
165 170 175	
ACC GTA GAT AAT ACC ATC GAT GCA AGA ATG GTC GAC GTA AAT GGA AAA	576
Thr Val Asp Asn Thr Ile Asp Ala Arg Met Val Asp Val Asn Gly Lys	
180 185 190	
ATT TTG GAT AGA AAA TTG ATG GGA GAA GAT CTC TAC TGG GCA ATA AAC	624
Ile Leu Asp Arg Lys Leu Met Gly Glu Asp Leu Tyr Trp Ala Ile Asn	
195 200 205	
GGA GGA GGA GGA GGG AGC TAC GGC GTC GTA TTG GCC TAC AAA ATA AAC	672
Gly Gly Gly Gly Gly Ser Tyr Gly Val Val Leu Ala Tyr Lys Ile Asn	
210 215 220	
CTT GTT GAA GTC CCA GAA AAC GTC ACC GTT TTC AGA ATC TCC CGG ACG	720
Leu Val Glu Val Pro Glu Asn Val Thr Val Phe Arg Ile Ser Arg Thr	
225 230 235 240	
TTA GAA CAA AAT GCG ACG GAT ATC ATT CAC CGG TGG CAA CAA GTT GCA	768
Leu Glu Gln Asn Ala Thr Asp Ile Ile His Arg Trp Gln Gln Val Ala	
245 250 255	
CCG AAG CTT CCC GAC GAG CTT TTC ATA AGA ACA GTC ATT GAC GTA GTA	816
Pro Lys Leu Pro Asp Glu Leu Phe Ile Arg Thr Val Ile Asp Val Val	
260 265 270	
AAC GGC ACT GTT TCA TCT CAA AAG ACC GTC AGG ACA ACA TTC ATA GCA	864
Asn Gly Thr Val Ser Ser Gln Lys Thr Val Arg Thr Thr Phe Ile Ala	
275 280 285	

ATG TTT CTA GGA GAC ACG ACA ACT CTA CTG TCG ATA TTA AAC CGG AGA	912
Met Phe Leu Gly Asp Thr Thr Leu Leu Ser Ile Leu Asn Arg Arg	
290 295 300	
TTC CCA GAA TTG GGT TTG GTC CGG TCT GAC TGT ACC GAA ACA AGC TGG	960
Phe Pro Glu Leu Gly Leu Val Arg Ser Asp Cys Thr Glu Thr Ser Trp	
305 310 315 320	
ATC CAA TCT GTG CTA TTC TGG ACA AAT ATC CAA GTT GGT TCG TCG GAG	1008
Ile Gln Ser Val Leu Phe Trp Thr Asn Ile Gln Val Gly Ser Ser Glu	
325 330 335	
ACA CTT CTA CTC CAA AGG AAT CAA CCC GTG AAC TAC CTC AAG AGG AAA	1056
Thr Leu Leu Leu Gln Arg Asn Gln Pro Val Asn Tyr Leu Lys Arg Lys	
340 345 350	
TCA GAT TAC GTA CGT GAA CCG ATT TCA AGA ACC GGT TTA GAG TCA ATT	1104
Ser Asp Tyr Val Arg Glu Pro Ile Ser Arg Thr Gly Leu Glu Ser Ile	
355 360 365	
TGG AAG AAA ATG ATC GAG CTT GAA ATT CCG ACA ATG GCT TTC AAT CCA	1152
Trp Lys Lys Met Ile Glu Leu Glu Ile Pro Thr Met Ala Phe Asn Pro	
370 375 380	
TAC GGT GGT GAG ATG GGG AGG ATA TCA TCT ACG GTG ACT CCG TTC CCA	1200
Tyr Gly Gly Glu Met Gly Arg Ile Ser Ser Thr Val Thr Pro Phe Pro	
385 390 395 400	
TAC AGA GCC GGT AAT CTC TGG AAG ATT CAG TAC GGT GCG AAT TGG AGA	1248
Tyr Arg Ala Gly Asn Leu Trp Lys Ile Gln Tyr Gly Ala Asn Trp Arg	
405 410 415	
GAT GAG ACT TTA ACC GAC CGG TAC ATG GAA TTG ACG AGG AAG TTG TAC	1296
Asp Glu Thr Leu Thr Asp Arg Tyr Met Glu Leu Thr Arg Lys Leu Tyr	
420 425 430	
CAA TTC ATG ACA CCA TTT GTT TCC AAG AAT CCG AGA CAA TCG TTT TTC	1344
Gln Phe Met Thr Pro Phe Val Ser Lys Asn Pro Arg Gln Ser Phe Phe	
435 440 445	
AAT TAC CGT GAT GTT GAT TTG GGT ATT AAT TCT CAT AAT GGT AAA ATC	1392
Asn Tyr Arg Asp Val Asp Leu Gly Ile Asn Ser His Asn Gly Lys Ile	
450 455 460	
AGT AGT TAT GTG GAA GGT AAA CGT TAC GGG AAG AAG TAT TTC GCA GGT	1440
Ser Ser Tyr Val Glu Gly Lys Arg Tyr Gly Lys Lys Tyr Phe Ala Gly	
465 470 475 480	
AAT TTC GAG AGA TTG GTG AAG ATT AAG ACG AGA GTT GAT AGT GGT AAT	1488
Asn Phe Glu Arg Leu Val Lys Ile Lys Thr Arg Val Asp Ser Gly Asn	
485 490 495	
TTC TTT AGG AAC GAA CAG AGT ATT CCT GTG TTA CCA TAA	1527
Phe Phe Arg Asn Glu Gln Ser Ile Pro Val Leu Pro	
500 505	

(2) INFORMATION FOR SEQ ID NO: 73:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 508 amino acids

(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

```
Thr Ser Arg Arg Asn Ser Glu Thr Phe Thr Gln Cys Leu Thr Ser Asn
 1              5              10              15
Ser Asp Pro Lys His Pro Ile Ser Pro Ala Ile Phe Phe Ser Gly Asn
          20              25              30
Gly Ser Tyr Ser Ser Val Leu Gln Ala Asn Ile Arg Asn Leu Arg Phe
          35              40              45
Asn Thr Thr Ser Thr Pro Lys Pro Phe Leu Ile Ile Ala Ala Thr His
          50              55              60
Glu Ser His Val Gln Ala Ala Ile Thr Cys Gly Lys Arg His Asn Leu
          65              70              75              80
Gln Met Lys Ile Arg Ser Gly Gly His Asp Tyr Asp Gly Leu Ser Tyr
          85              90              95
Val Thr Tyr Ser Gly Lys Pro Phe Phe Val Leu Asp Met Phe Asn Leu
          100              105              110
Arg Ser Val Asp Val Asp Val Ala Ser Lys Thr Ala Trp Val Gln Thr
          115              120              125
Gly Ala Ile Leu Gly Glu Val Tyr Tyr Tyr Ile Trp Glu Lys Ser Lys
          130              135              140
Thr Leu Ala Tyr Pro Ala Gly Ile Cys Pro Thr Val Gly Val Gly Gly
          145              150              155              160
His Ile Ser Gly Gly Gly Tyr Gly Asn Met Met Arg Lys Tyr Gly Leu
          165              170              175
Thr Val Asp Asn Thr Ile Asp Ala Arg Met Val Asp Val Asn Gly Lys
          180              185              190
Ile Leu Asp Arg Lys Leu Met Gly Glu Asp Leu Tyr Trp Ala Ile Asn
          195              200              205
Gly Gly Gly Gly Gly Ser Tyr Gly Val Val Leu Ala Tyr Lys Ile Asn
          210              215              220
Leu Val Glu Val Pro Glu Asn Val Thr Val Phe Arg Ile Ser Arg Thr
          225              230              235              240
Leu Glu Gln Asn Ala Thr Asp Ile Ile His Arg Trp Gln Gln Val Ala
          245              250              255
Pro Lys Leu Pro Asp Glu Leu Phe Ile Arg Thr Val Ile Asp Val Val
          260              265              270
Asn Gly Thr Val Ser Ser Gln Lys Thr Val Arg Thr Thr Phe Ile Ala
          275              280              285
```

Met Phe Leu Gly Asp Thr Thr Thr Leu Leu Ser Ile Leu Asn Arg Arg  
 290 395 300  
 Phe Pro Glu Leu Gly Leu Val Arg Ser Asp Cys Thr Glu Thr Ser Trp  
 305 310 315 320  
 Ile Gln Ser Val Leu Phe Trp Thr Asn Ile Gln Val Gly Ser Ser Glu  
 325 330 335  
 Thr Leu Leu Leu Gln Arg Asn Gln Pro Val Asn Tyr Leu Lys Arg Lys  
 340 345 350  
 Ser Asp Tyr Val Arg Glu Pro Ile Ser Arg Thr Gly Leu Glu Ser Ile  
 355 360 365  
 Trp Lys Lys Met Ile Glu Leu Glu Ile Pro Thr Met Ala Phe Asn Pro  
 370 375 380  
 Tyr Gly Gly Glu Met Gly Arg Ile Ser Ser Thr Val Thr Pro Phe Pro  
 385 390 395 400  
 Tyr Arg Ala Gly Asn Leu Trp Lys Ile Gln Tyr Gly Ala Asn Trp Arg  
 405 410 415  
 Asp Glu Thr Leu Thr Asp Arg Tyr Met Glu Leu Thr Arg Lys Leu Tyr  
 420 425 430  
 Gln Phe Met Thr Pro Phe Val Ser Lys Asn Pro Arg Gln Ser Phe Phe  
 435 440 445  
 Asn Tyr Arg Asp Val Asp Leu Gly Ile Asn Ser His Asn Gly Lys Ile  
 450 455 460  
 Ser Ser Tyr Val Glu Gly Lys Arg Tyr Gly Lys Lys Tyr Phe Ala Gly  
 465 470 475 480  
 Asn Phe Glu Arg Leu Val Lys Ile Lys Thr Arg Val Asp Ser Gly Asn  
 485 490 495  
 Phe Phe Arg Asn Glu Gln Ser Ile Pro Val Leu Pro  
 500 505

(2) INFORMATION FOR SEQ ID NO: 74:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1530 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: Colombia



(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..1527

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:

TCC ATC CAA GAT CAA TTC ATA AAC TGT GTC AAA AGA AAC ACA CAT GTT	48
Ser Ile Gln Asp Gln Phe Ile Asn Cys Val Lys Arg Asn Thr His Val	
1 5 10 15	
TCT TTT CCA CTC GAG AAA ACG TTA TTC ACC CCT GCG AAA AAC GTC TCT	96
Ser Phe Pro Leu Glu Lys Thr Leu Phe Thr Pro Ala Lys Asn Val Ser	
20 25 30	
TTG TTC AAC CAA GTC CTT GAA TCG ACG GCT CAA AAT CTC CAG TTC TTG	144
Leu Phe Asn Gln Val Leu Glu Ser Thr Ala Gln Asn Leu Gln Phe Leu	
35 40 45	
GCA AAA TCC ATG CCT AAA CCG GGA TTC ATA TTC AGA CCG ATT CAC CAG	192
Ala Lys Ser Met Pro Lys Pro Gly Phe Ile Phe Arg Pro Ile His Gln	
50 55 60	
TCT CAA GTC CAA GCT TCC ATC ATT TGT TCA AAG AAA CTC GGA ATT CAT	240
Ser Gln Val Gln Ala Ser Ile Ile Cys Ser Lys Lys Leu Gly Ile His	
65 70 75 80	
TTT CGT GTT AGA AGT GGC GGT CAC GAT TTC GAG GCC TTG TCT TAT GTT	288
Phe Arg Val Arg Ser Gly Gly His Asp Phe Glu Ala Leu Ser Tyr Val	
85 90 95	
TCA CGG ATT GAA AAA CCG TTT ATA TTA CTC GAC CTG TCA AAA TTG AAA	336
Ser Arg Ile Glu Lys Pro Phe Ile Leu Leu Asp Leu Ser Lys Leu Lys	
100 105 110	
CAA ATC AAT GTT GAT ATT GAA TCC AAT AGT GCT TGG GTT CAA CCT GGT	384
Gln Ile Asn Val Asp Ile Glu Ser Asn Ser Ala Trp Val Gln Pro Gly	
115 120 125	
GCT ACG CTT GGT GAG CTT TAC TAC AGA ATT GCA GAG AAG AGC AAG ATC	432
Ala Thr Leu Gly Glu Leu Tyr Tyr Arg Ile Ala Glu Lys Ser Lys Ile	
130 135 140	
CAT GGA TTT CCC GCG GGT TTG TGC ACA AGT GTA GGC ATA GGT GGG TAT	480
His Gly Phe Pro Ala Gly Leu Cys Thr Ser Val Gly Ile Gly Gly Tyr	
145 150 155 160	
ATG ACA GGC GGT GGA TAC GGT ACC TTG ATG AGG AAG TAT GGT CTT GCG	528
Met Thr Gly Gly Tyr Gly Thr Leu Met Arg Lys Tyr Gly Leu Ala	
165 170 175	
GGA GAT AAT GTT CTA GAC GTA AAG ATG GTT GAT GCA AAT GGT AAA TTA	576
Gly Asp Asn Val Leu Asp Val Lys Met Val Asp Ala Asn Gly Lys Leu	
180 185 190	
CTC GAC AGA GCC GCG ATG GGT GAG GAC CTA TTT TGG GCG ATT AGA GGA	624
Leu Asp Arg Ala Ala Met Gly Glu Asp Leu Phe Trp Ala Ile Arg Gly	
195 200 205	
GGC GGT GGA GCG AGT TTC GGG ATA GTT CTA GCA TGG AAG ATC AAG CTT	672
Gly Gly Gly Ala Ser Phe Gly Ile Val Leu Ala Trp Lys Ile Lys Leu	
210 215 220	

GTT	CCT	GTT	CCT	AAG	ACT	GTT	ACC	GTC	TTC	ACT	GTC	ACC	AAA	ACG	TTA	720
Val	Pro	Val	Pro	Lys	Thr	Val	Thr	Val	Phe	Thr	Val	Thr	Lys	Thr	Leu	
225					230					235					240	
GAA	CAA	GAC	GCA	AGA	TTG	AAG	ACT	ATT	TCT	AAG	TGG	CAA	CAA	ATT	TCA	768
Glu	Gln	Asp	Ala	Arg	Leu	Lys	Thr	Ile	Ser	Lys	Trp	Gln	Gln	Ile	Ser	
				245					250						255	
TCC	AAG	ATT	ATT	GAA	GAG	ATA	CAC	ATC	CGA	GTG	GTA	CTC	AGA	GCA	GCT	816
Ser	Lys	Ile	Ile	Glu	Glu	Ile	His	Ile	Arg	Val	Val	Leu	Arg	Ala	Ala	
			260					265							270	
GGA	AAT	GAT	GGA	AAC	AAG	ACT	GTG	ACA	ATG	ACC	TAC	CTA	GGT	CAG	TTT	864
Gly	Asn	Asp	Gly	Asn	Lys	Thr	Val	Thr	Met	Thr	Tyr	Leu	Gly	Gln	Phe	
		275					280								285	
CTT	GGC	GAG	AAA	GGC	ACC	TTG	CTG	AAG	GTT	ATG	GAG	AAG	GCT	TTT	CCA	912
Leu	Gly	Glu	Lys	Gly	Thr	Leu	Leu	Lys	Val	Met	Glu	Lys	Ala	Phe	Pro	
		290					295								300	
GAA	CTA	GGG	TTA	ACT	CAA	AAG	GAT	TGT	ACT	GAA	ATG	AGC	TGG	ATT	GAA	960
Glu	Leu	Gly	Leu	Thr	Gln	Lys	Asp	Cys	Thr	Glu	Met	Ser	Trp	Ile	Glu	
					310						315				320	
GCC	GCC	CTT	TTC	CAT	GGT	GGA	TTT	CCA	ACA	GGT	TCT	CCT	ATT	GAA	ATT	1008
Ala	Ala	Leu	Phe	His	Gly	Gly	Phe	Pro	Thr	Gly	Ser	Pro	Ile	Glu	Ile	
				325						330					335	
TTG	CTT	CAG	CTC	AAG	TCG	CCT	CTA	GGA	AAA	GAT	TAC	TTC	AAA	GCA	ACG	1056
Leu	Leu	Gln	Leu	Lys	Ser	Pro	Leu	Gly	Lys	Asp	Tyr	Phe	Lys	Ala	Thr	
			340						345						350	
TCG	GAT	TTC	GTT	AAA	GAA	CCT	ATT	CCT	GTG	ATA	GGC	TTC	AAA	GGA	ATA	1104
Ser	Asp	Phe	Val	Lys	Glu	Pro	Ile	Pro	Val	Ile	Gly	Phe	Lys	Gly	Ile	
		355						360							365	
TTC	AAA	AGA	TTG	ATT	GAA	GGA	AAC	ACA	ACA	TTT	CTG	AAC	TGG	ACT	CCT	1152
Phe	Lys	Arg	Leu	Ile	Glu	Gly	Asn	Thr	Thr	Phe	Leu	Asn	Trp	Thr	Pro	
		370					375								380	
TAC	GGT	GGT	ATG	ATG	TCG	AAA	ATC	CCT	GAA	TCT	GCG	ATC	CCA	TTT	CCG	1200
Tyr	Gly	Gly	Met	Met	Ser	Lys	Ile	Pro	Glu	Ser	Ala	Ile	Pro	Phe	Pro	
					390						395				400	
CAT	AGA	AAC	GGA	ACC	CTC	TTC	AAG	ATT	CTC	TAT	TAC	GCG	AAC	TGG	CTA	1248
His	Arg	Asn	Gly	Thr	Leu	Phe	Lys	Ile	Leu	Tyr	Tyr	Ala	Asn	Trp	Leu	
				405						410					415	
GAG	AAT	GAC	AAG	ACA	TCG	AGT	AGA	AAA	ATC	AAC	TGG	ATC	AAA	GAG	ATA	1296
Glu	Asn	Asp	Lys	Thr	Ser	Ser	Arg	Lys	Ile	Asn	Trp	Ile	Lys	Glu	Ile	
			420						425						430	
TAC	AAT	TAC	ATG	GCG	CCT	TAT	GTC	TCA	AGC	AAT	CCA	AGA	CAA	GCA	TAT	1344
Tyr	Asn	Tyr														

Val Asn Phe Ile Glu Ala Lys Ile Trp Gly Pro Lys Tyr Phe Lys Gly  
465 470 475 480

AAT TTT GAC AGA TTG GTG AAG ATT AAA ACC AAG GTT GAT CCA GAG AAC 1488  
Asn Phe Asp Arg Leu Val Lys Ile Lys Thr Lys Val Asp Pro Glu Asn  
485 490 495

TTC TTC AGG CAC GAG CAG AGT ATC CCA CCT ATG CCC TAC TAG 1530  
Phe Phe Arg His Glu Gln Ser Ile Pro Pro Met Pro Tyr  
500 505

(2) INFORMATION FOR SEQ ID NO: 75:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 509 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

Ser Ile Gln Asp Gln Phe Ile Asn Cys Val Lys Arg Asn Thr His Val  
1 5 10 15

Ser Phe Pro Leu Glu Lys Thr Leu Phe Thr Pro Ala Lys Asn Val Ser  
20 25 30

Leu Phe Asn Gln Val Leu Glu Ser Thr Ala Gln Asn Leu Gln Phe Leu  
35 40 45

Ala Lys Ser Met Pro Lys Pro Gly Phe Ile Phe Arg Pro Ile His Gln  
50 55 60

Ser Gln Val Gln Ala Ser Ile Ile Cys Ser Lys Lys Leu Gly Ile His  
65 70 75 80

Phe Arg Val Arg Ser Gly Gly His Asp Phe Glu Ala Leu Ser Tyr Val  
85 90 95

Ser Arg Ile Glu Lys Pro Phe Ile Leu Leu Asp Leu Ser Lys Leu Lys  
100 105 110

Gln Ile Asn Val Asp Ile Glu Ser Asn Ser Ala Trp Val Gln Pro Gly  
115 120 125

Ala Thr Leu Gly Glu Leu Tyr Tyr Arg Ile Ala Glu Lys Ser Lys Ile  
130 135 140

His Gly Phe Pro Ala Gly Leu Cys Thr Ser Val Gly Ile Gly Gly Tyr  
145 150 155 160

Met Thr Gly Gly Gly Tyr Gly Thr Leu Met Arg Lys Tyr Gly Leu Ala  
165 170 175

Gly Asp Asn Val Leu Asp Val Lys Met Val Asp Ala Asn Gly Lys Leu  
180 185 190

Leu Asp Arg Ala Ala Met Gly Glu Asp Leu Phe Trp Ala Ile Arg Gly  
195 200 205

Gly Gly Gly Ala Ser Phe Gly Ile Val Leu Ala Trp Lys Ile Lys Leu  
 210 215 220  
 Val Pro Val Pro Lys Thr Val Thr Val Phe Thr Val Thr Lys Thr Leu  
 225 230 235 240  
 Glu Gln Asp Ala Arg Leu Lys Thr Ile Ser Lys Trp Gln Gln Ile Ser  
 245 250 255  
 Ser Lys Ile Ile Glu Glu Ile His Ile Arg Val Val Leu Arg Ala Ala  
 260 265 270  
 Gly Asn Asp Gly Asn Lys Thr Val Thr Met Thr Tyr Leu Gly Gln Phe  
 275 280 285  
 Leu Gly Glu Lys Gly Thr Leu Leu Lys Val Met Glu Lys Ala Phe Pro  
 290 295 300  
 Glu Leu Gly Leu Thr Gln Lys Asp Cys Thr Glu Met Ser Trp Ile Glu  
 305 310 315 320  
 Ala Ala Leu Phe His Gly Gly Phe Pro Thr Gly Ser Pro Ile Glu Ile  
 325 330 335  
 Leu Leu Gln Leu Lys Ser Pro Leu Gly Lys Asp Tyr Phe Lys Ala Thr  
 340 345 350  
 Ser Asp Phe Val Lys Glu Pro Ile Pro Val Ile Gly Phe Lys Gly Ile  
 355 360 365  
 Phe Lys Arg Leu Ile Glu Gly Asn Thr Thr Phe Leu Asn Trp Thr Pro  
 370 375 380  
 Tyr Gly Gly Met Met Ser Lys Ile Pro Glu Ser Ala Ile Pro Phe Pro  
 385 390 395 400  
 His Arg Asn Gly Thr Leu Phe Lys Ile Leu Tyr Tyr Ala Asn Trp Leu  
 405 410 415  
 Glu Asn Asp Lys Thr Ser Ser Arg Lys Ile Asn Trp Ile Lys Glu Ile  
 420 425 430  
 Tyr Asn Tyr Met Ala Pro Tyr Val Ser Ser Asn Pro Arg Gln Ala Tyr  
 435 440 445  
 Val Asn Tyr Arg Asp Leu Asp Phe Gly Gln Asn Lys Asn Asn Ala Lys  
 450 455 460  
 Val Asn Phe Ile Glu Ala Lys Ile Trp Gly Pro Lys Tyr Phe Lys Gly  
 465 470 475 480  
 Asn Phe Asp Arg Leu Val Lys Ile Lys Thr Lys Val Asp Pro Glu Asn  
 485 490 495  
 Phe Phe Arg His Glu Gln Ser Ile Pro Pro Met Pro Tyr  
 500 505

**PATENT  
IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

In re: Maarten Hendrik STUIVER, et al

Serial No.: 09/258,031

Group No.: 1652

Filed: February 25, 1999

Examiner: R. Hutson

For: ANTIFUNGAL PROTEINS, DNA ENCODING THEREFOR, AND HOSTS  
INCORPORATING SAME

**Commissioner Patents and Trademarks  
Washington, DC 20231**

**AMENDMENT**

Sir:

In response to the Official Action of September 19, 1999, please amend the application  
as follows:

**IN THE SPECIFICATION:**

Page 51, line 1 - page 116, last line delete in entirety and substitute the following new

Sequence Listing:

---

**CERTIFICATE OF MAILING (37 CFR 1.8a)**

I hereby certify that this paper (along with any paper referred to as being attached or enclosed)  
is being deposited with the United States Postal on the date shown below with sufficient postage  
as first class mail in an envelope addressed to the: Commissioner of Patents and Trademarks,  
Washington, DC 20231

\_\_\_\_\_  
CLIFFORD J. MASS

Type or print name of person mailing paper)

Date: December 17, 1999

\_\_\_\_\_  
(Signature of person mailing paper)

## SEQUENCE LISTING

### (1) GENERAL INFORMATION:

(i) APPLICANT: STUIVER, Maarten Hendrik  
CUSTERS, Jerome Humbertina Henricus Victor  
SELA-BURLAGE, Marianne Beatrix  
MELCHERS, Leo Sjoerd  
VAN DEVENTER-TROOST, Johanna Pieterella  
LAGEWEG, Wessel  
PONSTEIN, Anne Silene  
LAGEWEG, Wessel  
PONSTEIN, Anne Silene

(ii) TITLE OF INVENTION: ANTIFUNGAL PROTEINS, DNA CODING  
THEREFOR, AND HOSTS INCORPORATING  
SAME.

(iii) NUMBER OF SEQUENCES: 75

#### (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: LADAS & PARRY  
(B) STREET: 26 WEST 61 STREET  
(C) CITY: NEW YORK  
(D) STATE: NY  
(E) COUNTRY: USA  
(F) ZIP: 10023 - 7604

#### (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3.25" Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: WINDOWS 95  
(D) SOFTWARE: WORDPERFECT 8

#### (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 09/258,031  
(B) FILING DATE: 25-FEB-1999  
(C) CLASSIFICATION: 435

#### (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PCT/EP97/04923  
(B) FILING DATE: 04-SEP-1997

#### (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: EP97200831.2  
(B) FILING DATE: 19-MAR-1997

#### (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: EP96202466.7  
(B) FILING DATE: 04-SEP-1996

### (2) INFORMATION FOR SEQ ID NO: 1:

#### (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helianthus annuus

(B) STRAIN: cv. zebulon

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Ser Ile Asn Val Asp Ile Glu Gln Glu Thr Ala Trp Val Gln Ala Gly  
1                      5                      10                      15

Ala Thr Leu Gly Glu Val Tyr Tyr Arg  
                    20                      25

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helianthus annuus

(B) STRAIN: cv. zebulon

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Asp Pro Ser Phe Pro Ile Thr Gly Glu Val Tyr Thr Pro Gly Xaa Ser  
1                      5                      10                      15

Ser Phe Pro Thr Val Leu Gln Asn Tyr  
                    20                      25

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: YES

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

- (B) LOCATION: 1
- (D) OTHER INFORMATION: /function= "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

AACTTCTCCN AGNGTNGCNC CNGCTTGNAC CCA

33

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 32 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: YES

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1
- (D) OTHER INFORMATION: /function= "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GATCCNTCTT TCCNATTAC TGGNGAGGTT TA

32

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 354 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Helianthus annuus
- (B) STRAIN: cv. zebulon

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..354

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GAT CCG TCT TTC CCG ATT ACT GGG GAG GTT TAC ACT CCC GGA AAC TCA  
 Asp Pro Ser Phe Pro Ile Thr Gly Glu Val Tyr Thr Pro Gly Asn Ser  
 1 5 10 15

48



TCT TTT CCT ACC GTC TTG CAA AAC TAC ATC CGA AAC CTT CGG TTC AAT	96
Ser Phe Pro Thr Val Leu Gln Asn Tyr Ile Arg Asn Leu Arg Phe Asn	
20 25 30	
GAA ACT ACC ACA CCA AAA CCC TTT TTA ATC ATC ACA GCC GAA CAT GTT	144
Glu Thr Thr Thr Pro Lys Pro Phe Leu Ile Ile Thr Ala Glu His Val	
35 40 45	
TCC CAC ATT CAG GCA GCT GTG GTT TGT GGC AAA CAA AAC CGG TTG CTA	192
Ser His Ile Gln Ala Ala Val Val Cys Gly Lys Gln Asn Arg Leu Leu	
50 55 60	
CTG AAA ACC AGA AGC GGT GGT CAT GAT TAT GAA GGT CTT TCC TAC CTT	240
Leu Lys Thr Arg Ser Gly Gly His Asp Tyr Glu Gly Leu Ser Tyr Leu	
65 70 75 80	
ACA AAC ACA AAC CAA CCC TTC TTC ATT GTG GAC ATG TTC AAT TTA AGG	288
Thr Asn Thr Asn Gln Pro Phe Phe Ile Val Asp Met Phe Asn Leu Arg	
85 90 95	
TCC ATA AAC GTA GAT ATC GAA CAA GAA ACC GCA TGG GTC CAA GCC GGC	336
Ser Ile Asn Val Asp Ile Glu Gln Glu Thr Ala Trp Val Gln Ala Gly	
100 105 110	
GCC ACC CTC GGA GAA GTT	354
Ala Thr Leu Gly Glu Val	
115	

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 118 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Asp Pro Ser Phe Pro Ile Thr Gly Glu Val Tyr Thr Pro Gly Asn Ser	
1 5 10 15	
Ser Phe Pro Thr Val Leu Gln Asn Tyr Ile Arg Asn Leu Arg Phe Asn	
20 25 30	
Glu Thr Thr Thr Pro Lys Pro Phe Leu Ile Ile Thr Ala Glu His Val	
35 40 45	
Ser His Ile Gln Ala Ala Val Val Cys Gly Lys Gln Asn Arg Leu Leu	
50 55 60	
Leu Lys Thr Arg Ser Gly Gly His Asp Tyr Glu Gly Leu Ser Tyr Leu	
65 70 75 80	
Thr Asn Thr Asn Gln Pro Phe Phe Ile Val Asp Met Phe Asn Leu Arg	
85 90 95	
Ser Ile Asn Val Asp Ile Glu Gln Glu Thr Ala Trp Val Gln Ala Gly	
100 105 110	

Ala Thr Leu Gly Glu Val  
115

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1
  - (D) OTHER INFORMATION: /function= "primer"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CAGGCAGCTG TGGTTTGTGG C

21

(2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1
  - (D) OTHER INFORMATION: /function= "primer"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GTCCACAATG AAGAAGGGTT G

21

(2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1

(D) OTHER INFORMATION: /function= "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

ACGTAGATAT CGAACAAGAA ACCGC

25

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GCTTTACTAC ACGGGCTTCC CCAG

24

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CTGGGGAAGC CCGTGTAGTA AAGC

24

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GGTACTCCAA CCACGGCGCT C

21

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

CGGGAAGTTG CAGAAGATTG GGTG

25

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GAGCAAGAGA AGAAGGAGAC

20

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1784 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helianthus annuus*

(B) STRAIN: Zebulon

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 21..1608

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

ATATCACATC TTCTTTCAAC ATG CAA ACT TCC ATT CTT ACT CTC CTT CTT	50
Met Gln Thr Ser Ile Leu Thr Leu Leu Leu	
1 5 10	
CTC TTG CTC TCA ACC CAA TCT TCT GCA ACT TCC CGT TCC ATT ACA GAT	98
Leu Leu Leu Ser Thr Gln Ser Ser Ala Thr Ser Arg Ser Ile Thr Asp	
15 20 25	
CGC TTC ATT CAA TGT TTA CAC GAC CGG GCC GAC CCT TCA TTT CCG ATA	146
Arg Phe Ile Gln Cys Leu His Asp Arg Ala Asp Pro Ser Phe Pro Ile	
30 35 40	
ACC GGA GAG GTT TAC ACT CCC GGA AAC TCA TCT TTT CCT ACC GTC TTG	194
Thr Gly Glu Val Tyr Thr Pro Gly Asn Ser Ser Phe Pro Thr Val Leu	
45 50 55	
CAA AAC TAC ATC CGA AAC CTT CGG TTC AAT GAA ACT ACC ACA CCA AAA	242
Gln Asn Tyr Ile Arg Asn Leu Arg Phe Asn Glu Thr Thr Thr Pro Lys	
60 65 70	
CCC TTT TTA ATC ATC ACA GCC GAA CAT GTT TCC CAC ATT CAG GCA GCT	290
Pro Phe Leu Ile Ile Thr Ala Glu His Val Ser His Ile Gln Ala Ala	
75 80 85 90	
GTG GTT TGT GGC AAA CAA AAC CGG TTG CTA CTG AAA ACC AGA AGC GGT	338
Val Val Cys Gly Lys Gln Asn Arg Leu Leu Leu Lys Thr Arg Ser Gly	
95 100 105	
GGT CAT GAT TAT GAA GGT CTT TCC TAC CTT ACA AAC ACA AAC CAA CCC	386
Gly His Asp Tyr Glu Gly Leu Ser Tyr Leu Thr Asn Thr Asn Gln Pro	
110 115 120	
TTC TTC ATT GTG GAC ATG TTC AAT TTA AGG TCC ATA AAC GTA GAT ATC	434
Phe Phe Ile Val Asp Met Phe Asn Leu Arg Ser Ile Asn Val Asp Ile	
125 130 135	
GAA CAA GAA ACC GCA TGG GTC CAA GCC GGT GCG ACT CTT GGT GAA GTG	482
Glu Gln Glu Thr Ala Trp Val Gln Ala Gly Ala Thr Leu Gly Glu Val	
140 145 150	
TAC TAT CGA ATA GCG GAG AAA AGT AAC AAG CAT GGT TTT CCG GCA GGG	530
Tyr Tyr Arg Ile Ala Glu Lys Ser Asn Lys His Gly Phe Pro Ala Gly	
155 160 165 170	
GTT TGT CCA ACG GTT GGC GTT GGT GGG CAT TTT AGT GGT GGT GGG TAT	578
Val Cys Pro Thr Val Gly Val Gly Gly His Phe Ser Gly Gly Gly Tyr	
175 180 185	
GGT AAT TTG ATG AGA AAA TAT GGT TTG TCG GTT GAT AAT ATT GTT GAT	626
Gly Asn Leu Met Arg Lys Tyr Gly Leu Ser Val Asp Asn Ile Val Asp	
190 195 200	

GCT CAA ATA ATA GAT GTG AAT GGC AAG CTT TTG GAT CGA AAG AGT ATG	674
Ala Gln Ile Ile Asp Val Asn Gly Lys Leu Leu Asp Arg Lys Ser Met	
205 210 215	
GGT GAG GAT TTG TTT TGG GCG ATC ACC GGC GGT GGT GGT GTT AGT TTT	722
Gly Glu Asp Leu Phe Trp Ala Ile Thr Gly Gly Gly Gly Val Ser Phe	
220 225 230	
GGT GTG GTT CTA GCC TAC AAA ATC AAA CTA GTT CGT GTT CCG GAG GTT	770
Gly Val Val Leu Ala Tyr Lys Ile Lys Leu Val Arg Val Pro Glu Val	
235 240 245 250	
GTG ACC GTG TTT ACC ATT GAA AGA AGA GAG GAA CAA AAC CTC AGC ACC	818
Val Thr Val Phe Thr Ile Glu Arg Arg Glu Glu Gln Asn Leu Ser Thr	
255 260 265	
ATC GCG GAA CGA TGG GTA CAA GTT GCT GAT AAG CTA GAT AGA GAT CTT	866
Ile Ala Glu Arg Trp Val Gln Val Ala Asp Lys Leu Asp Arg Asp Leu	
270 275 280	
TTC CTT CGA ATG ACC TTT AGT GTC ATA AAC GAT ACC AAC GGT GGA AAG	914
Phe Leu Arg Met Thr Phe Ser Val Ile Asn Asp Thr Asn Gly Gly Lys	
285 290 295	
ACA GTC CGT GCT ATC TTT CCA ACG TTG TAC CTT GGA AAC TCG AGG AAT	962
Thr Val Arg Ala Ile Phe Pro Thr Leu Tyr Leu Gly Asn Ser Arg Asn	
300 305 310	
CTT GTT ACA CTT TTG AAT AAA GAT TTC CCC GAG TTA GGG TTG CAA GAA	1010
Leu Val Thr Leu Leu Asn Lys Asp Phe Pro Glu Leu Gly Leu Gln Glu	
315 320 325 330	
TCG GAT TGT ACT GAA ATG AGT TGG GTT GAG TCT GTG CTT TAC TAC ACG	1058
Ser Asp Cys Thr Glu Met Ser Trp Val Glu Ser Val Leu Tyr Tyr Thr	
335 340 345	
GGC TTC CCC AGT GGT ACT CCA ACC ACG GCG CTC TTA AGC CGT ACT CCT	1106
Gly Phe Pro Ser Gly Thr Pro Thr Thr Ala Leu Leu Ser Arg Thr Pro	
350 355 360	
CAA AGA CTC AAC CCA TTC AAG ATC AAA TCC GAT TAT GTG CAA AAT CCT	1154
Gln Arg Leu Asn Pro Phe Lys Ile Lys Ser Asp Tyr Val Gln Asn Pro	
365 370 375	
ATT TCT AAA CGA CAG TTC GAG TTC ATC TTC GAA AGG CTG AAA GAA CTT	1202
Ile Ser Lys Arg Gln Phe Glu Phe Ile Phe Glu Arg Leu Lys Glu Leu	
380 385 390	
GAA AAC CAA ATG TTG GCT TTC AAC CCA TAT GGT GGT AGA ATG AGT GAA	1250
Glu Asn Gln Met Leu Ala Phe Asn Pro Tyr Gly Gly Arg Met Ser Glu	
395 400 405 410	
ATA TCC GAA TTC GCA AAG CCT TTC CCA CAT AGA TCG GGT AAC ATA GCG	1298
Ile Ser Glu Phe Ala Lys Pro Phe Pro His Arg Ser Gly Asn Ile Ala	
415 420 425	
AAA ATT CAA TAC GAA GTA AAC TGG GAG GAT CTT AGC GAT GAA GCC GAA	1346
Lys Ile Gln Tyr Glu Val Asn Trp Glu Asp Leu Ser Asp Glu Ala Glu	
430 435 440	

AAT CGT TAC TTG AAT TTC ACA AGG CTG ATG TAT GAT TAC ATG ACC CCA 1394  
 Asn Arg Tyr Leu Asn Phe Thr Arg Leu Met Tyr Asp Tyr Met Thr Pro  
 445 450 455  
  
 TTT GTG TCG AAA AAC CCT AGA AAA GCA TTT TTG AAC TAT AGG GAT TTG 1442  
 Phe Val Ser Lys Asn Pro Arg Lys Ala Phe Leu Asn Tyr Arg Asp Leu  
 460 465 470  
  
 GAT ATT GGT ATC AAC AGC CAT GGC AGG AAT GCT TAT ACT GAA GGA ATG 1490  
 Asp Ile Gly Ile Asn Ser His Gly Arg Asn Ala Tyr Thr Glu Gly Met  
 475 480 485 490  
  
 GTT TAT GGG CAC AAG TAT TTC AAA GAG ACA AAT TAC AAG AGG CTA GTA 1538  
 Val Tyr Gly His Lys Tyr Phe Lys Glu Thr Asn Tyr Lys Arg Leu Val  
 495 500 505  
  
 AGT GTG AAG ACT AAA GTT GAT CCT GAC AAC TTC TTT AGG AAT GAG CAA 1586  
 Ser Val Lys Thr Lys Val Asp Pro Asp Asn Phe Phe Arg Asn Glu Gln  
 510 515 520  
  
 AGC ATC CCA ACT TTG TCA TCT T GAAGAACGTA CATATATAAA TAAATACCTT 1638  
 Ser Ile Pro Thr Leu Ser Ser  
 525  
  
 TGTGCATGGT ATTTTCAGGG TGTTAAAGTG ATATTCAGAT ATTTATGATA GAATTTTGAC 1698  
 TTGTATTTTA TACAATCAAA ATTGTATGGT TCTCCGAATT TCTCTTTTTA ATTCTGAAAA 1758  
 ATACATATTA GTATTGTCAA AAAAAA 1784

(2) INFORMATION FOR SEQ ID NO: 16:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 529 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Met Gln Thr Ser Ile Leu Thr Leu Leu Leu Leu Leu Ser Thr Gln  
 1 5 10 15  
 Ser Ser Ala Thr Ser Arg Ser Ile Thr Asp Arg Phe Ile Gln Cys Leu  
 20 25 30  
 His Asp Arg Ala Asp Pro Ser Phe Pro Ile Thr Gly Glu Val Tyr Thr  
 35 40 45  
 Pro Gly Asn Ser Ser Phe Pro Thr Val Leu Gln Asn Tyr Ile Arg Asn  
 50 55 60  
 Leu Arg Phe Asn Glu Thr Thr Thr Pro Lys Pro Phe Leu Ile Ile Thr  
 65 70 75 80  
 Ala Glu His Val Ser His Ile Gln Ala Ala Val Val Cys Gly Lys Gln  
 85 90 95

Asn Arg Leu Leu Leu Lys Thr Arg Ser Gly Gly His Asp Tyr Glu Gly  
 100 105 110  
 Leu Ser Tyr Leu Thr Asn Thr Asn Gln Pro Phe Phe Ile Val Asp Met  
 115 120 125  
 Phe Asn Leu Arg Ser Ile Asn Val Asp Ile Glu Gln Glu Thr Ala Trp  
 130 135 140  
 Val Gln Ala Gly Ala Thr Leu Gly Glu Val Tyr Tyr Arg Ile Ala Glu  
 145 150 155 160  
 Lys Ser Asn Lys His Gly Phe Pro Ala Gly Val Cys Pro Thr Val Gly  
 165 170 175  
 Val Gly Gly His Phe Ser Gly Gly Gly Tyr Gly Asn Leu Met Arg Lys  
 180 185 190  
 Tyr Gly Leu Ser Val Asp Asn Ile Val Asp Ala Gln Ile Ile Asp Val  
 195 200 205  
 Asn Gly Lys Leu Leu Asp Arg Lys Ser Met Gly Glu Asp Leu Phe Trp  
 210 215 220  
 Ala Ile Thr Gly Gly Gly Gly Val Ser Phe Gly Val Val Leu Ala Tyr  
 225 230 235 240  
 Lys Ile Lys Leu Val Arg Val Pro Glu Val Val Thr Val Phe Thr Ile  
 245 250 255  
 Glu Arg Arg Glu Glu Gln Asn Leu Ser Thr Ile Ala Glu Arg Trp Val  
 260 265 270  
 Gln Val Ala Asp Lys Leu Asp Arg Asp Leu Phe Leu Arg Met Thr Phe  
 275 280 285  
 Ser Val Ile Asn Asp Thr Asn Gly Gly Lys Thr Val Arg Ala Ile Phe  
 290 295 300  
 Pro Thr Leu Tyr Leu Gly Asn Ser Arg Asn Leu Val Thr Leu Leu Asn  
 305 310 315 320  
 Lys Asp Phe Pro Glu Leu Gly Leu Gln Glu Ser Asp Cys Thr Glu Met  
 325 330 335  
 Ser Trp Val Glu Ser Val Leu Tyr Tyr Thr Gly Phe Pro Ser Gly Thr  
 340 345 350  
 Pro Thr Thr Ala Leu Leu Ser Arg Thr Pro Gln Arg Leu Asn Pro Phe  
 355 360 365  
 Lys Ile Lys Ser Asp Tyr Val Gln Asn Pro Ile Ser Lys Arg Gln Phe  
 370 375 380  
 Glu Phe Ile Phe Glu Arg Leu Lys Glu Leu Glu Asn Gln Met Leu Ala  
 385 390 395 400  
 Phe Asn Pro Tyr Gly Gly Arg Met Ser Glu Ile Ser Glu Phe Ala Lys  
 405 410 415



[illegible]

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

CCGCCATGGA GACTTCCATT CTTACTC

27

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GCCGGATCCT CAAGATGACA AAGTTGGGAT GCT

33

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1589 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helianthus annuus*
- (B) STRAIN: Zebulon

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1590

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

ATG GAG ACT TCC ATT CTT ACT CTC CTT CTT CTC TTG CTC TCA ACC CAA	48
Met Glu Thr Ser Ile Leu Thr Leu Leu Leu Leu Leu Ser Thr Gln	
1 5 10 15	
TCT TCT GCA ACT TCC CGT TCC ATT ACA GAT CGC TTC ATT CAA TGT TTA	96
Ser Ser Ala Thr Ser Arg Ser Ile Thr Asp Arg Phe Ile Gln Cys Leu	
20 25 30	
CAC GAC CGG GCC GAC CCT TCA TTT CCG ATA ACC GGA GAG GTT TAC ACT	144
His Asp Arg Ala Asp Pro Ser Phe Pro Ile Thr Gly Glu Val Tyr Thr	
35 40 45	
CCC GGA AAC TCA TCT TTT CCT ACC GTC TTG CAA AAC TAC ATC CGA AAC	192
Pro Gly Asn Ser Ser Phe Pro Thr Val Leu Gln Asn Tyr Ile Arg Asn	
50 55 60	
CTT CGG TTC AAT GAA ACT ACC ACA CCA AAA CCC TTT TTA ATC ATC ACA	240
Leu Arg Phe Asn Glu Thr Thr Thr Pro Lys Pro Phe Leu Ile Ile Thr	
65 70 75 80	
GCC GAA CAT GTT TCC CAC ATT CAG GCA GCT GTG GTT TGT GGC AAA CAA	288
Ala Glu His Val Ser His Ile Gln Ala Ala Val Val Cys Gly Lys Gln	
85 90 95	
AAC CGG TTG CTA CTG AAA ACC AGA AGC GGT GGT CAT GAT TAT GAA GGT	336
Asn Arg Leu Leu Leu Lys Thr Arg Ser Gly Gly His Asp Tyr Glu Gly	
100 105 110	
CTT TCC TAC CTT ACA AAC ACA AAC CAA CCC TTC TTC ATT GTG GAC ATG	384
Leu Ser Tyr Leu Thr Asn Thr Asn Gln Pro Phe Phe Ile Val Asp Met	
115 120 125	
TTC AAT TTA AGG TCC ATA AAC GTA GAT ATC GAA CAA GAA ACC GCA TGG	432
Phe Asn Leu Arg Ser Ile Asn Val Asp Ile Glu Gln Glu Thr Ala Trp	
130 135 140	

GTC CAA GCC GGT GCG ACT CTT GGT GAA GTG TAC TAT CGA ATA GCG GAG Val Gln Ala Gly Ala Thr Leu Gly Glu Val Tyr Tyr Arg Ile Ala Glu 145 150 155 160	480
AAA AGT AAC AAG CAT GGT TTT CCG GCA GGG GTT TGT CCA ACG GTT GGC Lys Ser Asn Lys His Gly Phe Pro Ala Gly Val Cys Pro Thr Val Gly 165 170 175	528
GTT GGT GGG CAT TTT AGT GGT GGT GGG TAT GGT AAT TTG ATG AGA AAA Val Gly Gly His Phe Ser Gly Gly Gly Tyr Gly Asn Leu Met Arg Lys 180 185 190	576
TAT GGT TTG TCG GTT GAT AAT ATT GTT GAT GCT CAA ATA ATA GAT GTG Tyr Gly Leu Ser Val Asp Asn Ile Val Asp Ala Gln Ile Ile Asp Val 195 200 205	624
AAT GGC AAG CTT TTG GAT CGA AAG AGT ATG GGT GAG GAT TTG TTT TGG Asn Gly Lys Leu Leu Asp Arg Lys Ser Met Gly Glu Asp Leu Phe Trp 210 215 220	672
GCG ATC ACC GGC GGT GGT GGT GTT AGT TTT GGT GTG GTT CTA GCC TAC Ala Ile Thr Gly Gly Gly Gly Val Ser Phe Gly Val Val Leu Ala Tyr 225 230 235 240	720
AAA ATC AAA CTA GTT CGT GTT CCG GAG GTT GTG ACC GTG TTT ACC ATT Lys Ile Lys Leu Val Arg Val Pro Glu Val Val Thr Val Phe Thr Ile 245 250 255	768
GAA AGA AGA GAG GAA CAA AAC CTC AGC ACC ATC GCG GAA CGA TGG GTA Glu Arg Arg Glu Glu Gln Asn Leu Ser Thr Ile Ala Glu Arg Trp Val 260 265 270	816
CAA GTT GCT GAT AAG CTA GAT AGA GAT CTT TTC CTT CGA ATG ACC TTT Gln Val Ala Asp Lys Leu Asp Arg Asp Leu Phe Leu Arg Met Thr Phe 275 280 285	864
AGT GTC ATA AAC GAT ACC AAC GGT GGA AAG ACA GTC CGT GCT ATC TTT Ser Val Ile Asn Asp Thr Asn Gly Gly Lys Thr Val Arg Ala Ile Phe 290 295 300	912
CCA ACG TTG TAC CTT GGA AAC TCG AGG AAT CTT GTT ACA CTT TTG AAT Pro Thr Leu Tyr Leu Gly Asn Ser Arg Asn Leu Val Thr Leu Leu Asn 305 310 315 320	960
AAA GAT TTC CCC GAG TTA GGG TTG CAA GAA TCG GAT TGT ACT GAA ATG Lys Asp Phe Pro Glu Leu Gly Leu Gln Glu Ser Asp Cys Thr Glu Met 325 330 335	1008
AGT TGG GTT GAG TCT GTG CTT TAC TAC ACG GGC TTC CCC AGT GGT ACT Ser Trp Val Glu Ser Val Leu Tyr Tyr Thr Gly Phe Pro Ser Gly Thr 340 345 350	1056
CCA ACC ACG GCG CTC TTA AGC CGT ACT CCT CAA AGA CTC AAC CCA TTC Pro Thr Thr Ala Leu Leu Ser Arg Thr Pro Gln Arg Leu Asn Pro Phe 355 360 365	1104
AAG ATC AAA TCC GAT TAT GTG CAA AAT CCT ATT TCT AAA CGA CAG TTC Lys Ile Lys Ser Asp Tyr Val Gln Asn Pro Ile Ser Lys Arg Gln Phe 370 375 380	1152

GAG TTC ATC TTC GAA AGG ATG AAA GAA CTT GAA AAC CAA ATG TTG GCG	1200
Glu Phe Ile Phe Glu Arg Met Lys Glu Leu Glu Asn Gln Met Leu Ala	
385 390 395 400	
TTC AAC CCA TAT GGT GGT AGA ATG AGT GAA ATA TCC GAA TTC GCA AAG	1248
Phe Asn Pro Tyr Gly Gly Arg Met Ser Glu Ile Ser Glu Phe Ala Lys	
405 410 415	
CCT TTC CCA CAT AGA TCG GGT AAC ATA GCG AAG ATT CAA TAC GAA GTA	1296
Pro Phe Pro His Arg Ser Gly Asn Ile Ala Lys Ile Gln Tyr Glu Val	
420 425 430	
AAC TGG GAG GAT CTT AGC GAT GAA GCC GAA AAT CGT TAC TTG AAT TTC	1344
Asn Trp Glu Asp Leu Ser Asp Glu Ala Glu Asn Arg Tyr Leu Asn Phe	
435 440 445	
ACA AGG CTG ATG TAT GAT TAC ATG ACT CCA TTT GTG TCG AAA AAC CCT	1392
Thr Arg Leu Met Tyr Asp Tyr Met Thr Pro Phe Val Ser Lys Asn Pro	
450 455 460	
AGA GAA GCA TTT TTG AAC TAT AGG GAT TTG GAT ATT GGT ATC AAC AGC	1440
Arg Glu Ala Phe Leu Asn Tyr Arg Asp Leu Asp Ile Gly Ile Asn Ser	
465 470 475 480	
CAT GGC AGG AAT GCT TAT ACT GAA GGA ATG GTT TAT GGG CAC AAA TAT	1488
His Gly Arg Asn Ala Tyr Thr Glu Gly Met Val Tyr Gly His Lys Tyr	
485 490 495	
TTC AAA GAG ACA AAT TAC AAG AGG CTA GTA AGT GTG AAG ACT AAA GTT	1536
Phe Lys Glu Thr Asn Tyr Lys Arg Leu Val Ser Val Lys Thr Lys Val	
500 505 510	
GAT CCT GAC AAC TTC TTT AGG AAT GAG CAA AGC ATC CCA ACT TTG TCA	1584
Asp Pro Asp Asn Phe Phe Arg Asn Glu Gln Ser Ile Pro Thr Leu Ser	
515 520 525	
TCT TG	1589
Ser	
530	

(2) INFORMATION FOR SEQ ID NO: 20:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 529 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

Met Glu Thr Ser Ile Leu Thr Leu Leu Leu Leu Leu Ser Thr Gln	
1 5 10 15	
Ser Ser Ala Thr Ser Arg Ser Ile Thr Asp Arg Phe Ile Gln Cys Leu	
20 25 30	
His Asp Arg Ala Asp Pro Ser Phe Pro Ile Thr Gly Glu Val Tyr Thr	
35 40 45	

Pro Gly Asn Ser Ser Phe Pro Thr Val Leu Gln Asn Tyr Ile Arg Asn  
 50 55 60

Leu Arg Phe Asn Glu Thr Thr Thr Pro Lys Pro Phe Leu Ile Ile Thr  
 65 70 75 80

Ala Glu His Val Ser His Ile Gln Ala Ala Val Val Cys Gly Lys Gln  
 85 90 95

Asn Arg Leu Leu Leu Lys Thr Arg Ser Gly Gly His Asp Tyr Glu Gly  
 100 105 110

Leu Ser Tyr Leu Thr Asn Thr Asn Gln Pro Phe Phe Ile Val Asp Met  
 115 120 125

Phe Asn Leu Arg Ser Ile Asn Val Asp Ile Glu Gln Glu Thr Ala Trp  
 130 135 140

Val Gln Ala Gly Ala Thr Leu Gly Glu Val Tyr Tyr Arg Ile Ala Glu  
 145 150 155 160

Lys Ser Asn Lys His Gly Phe Pro Ala Gly Val Cys Pro Thr Val Gly  
 165 170 175

Val Gly Gly His Phe Ser Gly Gly Gly Tyr Gly Asn Leu Met Arg Lys  
 180 185 190

Tyr Gly Leu Ser Val Asp Asn Ile Val Asp Ala Gln Ile Ile Asp Val  
 195 200 205

Asn Gly Lys Leu Leu Asp Arg Lys Ser Met Gly Glu Asp Leu Phe Trp  
 210 215 220

Ala Ile Thr Gly Gly Gly Gly Val Ser Phe Gly Val Val Leu Ala Tyr  
 225 230 235 240

Lys Ile Lys Leu Val Arg Val Pro Glu Val Val Thr Val Phe Thr Ile  
 245 250 255

Glu Arg Arg Glu Glu Gln Asn Leu Ser Thr Ile Ala Glu Arg Trp Val  
 260 265 270

Gln Val Ala Asp Lys Leu Asp Arg Asp Leu Phe Leu Arg Met Thr Phe  
 275 280 285

Ser Val Ile Asn Asp Thr Asn Gly Gly Lys Thr Val Arg Ala Ile Phe  
 290 295 300

Pro Thr Leu Tyr Leu Gly Asn Ser Arg Asn Leu Val Thr Leu Leu Asn  
 305 310 315 320

Lys Asp Phe Pro Glu Leu Gly Leu Gln Glu Ser Asp Cys Thr Glu Met  
 325 330 335

Ser Trp Val Glu Ser Val Leu Tyr Tyr Thr Gly Phe Pro Ser Gly Thr  
 340 345 350

Pro Thr Thr Ala Leu Leu Ser Arg Thr Pro Gln Arg Leu Asn Pro Phe  
 355 360 365

Lys Ile Lys Ser Asp Tyr Val Gln Asn Pro Ile Ser Lys Arg Gln Phe  
 370 375 380  
 Glu Phe Ile Phe Glu Arg Met Lys Glu Leu Glu Asn Gln Met Leu Ala  
 385 390 395 400  
 Phe Asn Pro Tyr Gly Gly Arg Met Ser Glu Ile Ser Glu Phe Ala Lys  
 405 410 415  
 Pro Phe Pro His Arg Ser Gly Asn Ile Ala Lys Ile Gln Tyr Glu Val  
 420 425 430  
 Asn Trp Glu Asp Leu Ser Asp Glu Ala Glu Asn Arg Tyr Leu Asn Phe  
 435 440 445  
 Thr Arg Leu Met Tyr Asp Tyr Met Thr Pro Phe Val Ser Lys Asn Pro  
 450 455 460  
 Arg Glu Ala Phe Leu Asn Tyr Arg Asp Leu Asp Ile Gly Ile Asn Ser  
 465 470 475 480  
 His Gly Arg Asn Ala Tyr Thr Glu Gly Met Val Tyr Gly His Lys Tyr  
 485 490 495  
 Phe Lys Glu Thr Asn Tyr Lys Arg Leu Val Ser Val Lys Thr Lys Val  
 500 505 510  
 Asp Pro Asp Asn Phe Phe Arg Asn Glu Gln Ser Ile Pro Thr Leu Ser  
 515 520 525  
 Ser

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 350 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..350

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GAGAAACTCG GAGACTTTCA CACAATGCCT AACCTCAAAC TCCGACCCCA AACATCCCAT 60  
 CTCCCCGCT ATCTTCTTCT CCGGAAATGG CTCCTACTCC TCCGTATTAC AAGCCAACAT 120  
 CCGTAACCTC CGCTTCAACA CCACCTCAAC TCCGAAACCC TTCCTCATAA TCGCCGCAAC 180

ACATGAATCC CATGTGCAAG CCGCGATTAC TTGCGGGAAA CGCCACAACC TTCAGATGAA 240  
 AATCAGAAGT GGAGGCCACG ACTACGATGG CTTGTCATAC GTTACATACT CTGGCAAACC 300  
 GTTCTTCGTC CTCGACATGT TTAACCTCCG TTCGGTGGAT GTCGACGTGG 350

(2) INFORMATION FOR SEQ ID NO: 22:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 278 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Arabidopsis thaliana
  - (B) STRAIN: ecotype Columbia
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 2..278
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

GGCATGGATC TCCGCCGAG CGACTCTCGG AGAGGTTTAT TATCGGATTT GGGAGAAAAG 60  
 CAGAGTCCAT GGATTCCCCG CCGGAGTTTG ACCGACGGTT GGTGTTGGTG GGCATTTAAG 120  
 CGGCGGTGGT TACG GTAACA TGGTGAGGAA GTTTGGATTA TCTGTGGATT ACGTTGAGGA 180  
 TGCCAAGATC GTCGATGTAA ACNGTCGGGT TTTAGATCGG AAAGCAATGG GTGAGGATCT 240  
 GTTCTGGGCG ATTACCGGTG GAGGAGGAGG TAGCGTAC 278

(2) INFORMATION FOR SEQ ID NO: 23:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 345 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Arabidopsis thaliana
  - (B) STRAIN: ecotype Columbia
- (ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 2..345

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

TGGACATATT AGCGGAGGAG GATTCGGTAC AATAATGAGG AAATACGGTT TAGCGTCTGA 60  
TAACGTTGTG GACGCACGTT TGATGGATGT AAATGGGAAA ACTCTTGACC GGAAAACGAT 120  
GGGAGAGGAT TTGTTTTGGG CGCTTAGAGG CGGTGGAGCT GCGAGTTTTG GCGTTGTCTT 180  
GTCGTGGAAG GTTAAGCTTG CTAGGGTTCC TGAAAAGGTA ACTTGTTTCA TAAGTCAACA 240  
TCCGATGGGA CCTAGCATGA ACAAGCTTGT TCATAGATGG CAATCCATAG GATCAAGANN 300  
GCTAGACGAA GATTTATTCA TCAGAGTCAA TATTGACAAC AGTCT 345

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 695 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana  
(B) STRAIN: ecotype Columbia

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 1..695

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

GTTCGTTAAA ACCTATCCTN NANGGGCNAA AGNATATCAA AGNTTGNTTA NGNAACCCAA 60  
NATTTCTGAA CTGGCCNCCT TCGGTGGTAT ATGNCNAAAN CCCTTGAATC TGCGNANCCN 120  
ATTCCGCATA GAAACGGAAC CCTCTTCAAG ATTCTCTATT TACNCGAACT GNCTAGANNG 180  
AATGACAAGA CATCGAGTAG NAAAATCAAC TGGATCAAAG AGATATACAA TTACATGGCG 240  
CCTTATGTCT CAAGCAATCC AAGACAAGCA TATGTGAACT ACAGAGATCT AGACTTCGGA 300  
CAGAACAAGA ACAACGCAAA GGTAACTTC ATTGAAGCTA AAATCTGGGG ACCTAAGTAC 360  
TTCAAAGGCA ATTTTGACAG ATTGGTGAAG ATTTAAACCA AGGTTGATCC AGAGAACTTC 420  
TTCAGGCACG AGCAGAGTAT CCCACCTATG CCCTACTAGA AGCTAGGTTC ATGAAACCAA 480  
TAACATTATC AAAAATAAGR ATAAATGRTA ATTGTATACA ACATGATTCTG KCTTTCTTTA 540  
TTTCAGACAA TGTGGGAACT ACTCTAAANT AAAAWGTCNA TTTACCTTAA AAAAAAATA 600



ATCCCCNNTA ANANAAAANT GGGGGGGCCN TTTTGGGGN TCCCGGTTTT NGGACGGGGN 660  
 GCTTTNGGGG GGCTTGGNNT TTTTTTNGGN GCCCC 695

(2) INFORMATION FOR SEQ ID NO: 25:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 495 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Arabidopsis thaliana
  - (B) STRAIN: ecotype Columbia
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 2..495
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

TCTGTTTNA GGCAGAGCAG AGGAAGTTGT TGCTTTGCTT GGTAAGGAGT TTCCTGAATT 60  
 NAGTTTAAAG AAGGAGAACT GTTCGGAGAT GACTTGGTTT CAGTCAGCTT TATGGTGGGA 120  
 TAATCGTGTT AACCTACTC ANATTGATCC WAAAGTGTTT CTCGATCGGA ATCTTGATAG 180  
 AGCGAATTC GGAAAGAGGA AATCGGATTA CGTTGCGAGT AAGATTCCTA GAGATGGGAT 240  
 TAAGYCTTTT TCCAAGARGA TGMCTGACCT GGGGAAAAYC GGGCTTGTTT TTAAWCCGTA 300  
 TGGTGGGAAA ATGGCGGAGG TTACGGTTAA CGCGACGCCG TTTCCNCACC GAAGCAAGCT 360  
 TTTTAAGATT CAGTACTCGG TGACTIONGCA AGAAAACCTCT NTCGAGATAG AGAAAGGGTT 420  
 TCTTGAATCA GGCTAACGTC CTTATAGGTT CATGACCGGG TTTTINAGCA AGANCCCTGG 480  
 AATNCTTACT TNAAT 495

(2) INFORMATION FOR SEQ ID NO: 26:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 204 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..204

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

```
AAATTAAAAC AAATCAATGT TGATATTGAA TCCAATAGTG CTTGGTTTCA ACCTGGTGCT 60
ACGCTTGGTG AGCTTTACTA CAGAATTNCA GAGAAGAGCA AAATCCATGG ATTTCCNGCG 120
GGTTTNTNCA CAAGCNTAGG CATAGGTGGG TATATNANAG GCGGTGGATA CGGTACCTTG 180
ATGAGGAAGT ATGGTCTTNC GGGA                                     204
```

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 491 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..491

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

```
GAGATTTCTC GAGCAAGATA CTCCACTGAT GATCTTTGAG CCATTGGGTG GGAAAATCAG 60
CAAGATTTCA GAAACAGAAT CTCCATATCC ACACAGAAGA GGTAATCTGT ATAATATACA 120
GTACATGGTG AAATGGAAAG TGAATGANGT CGAGGAGATG AACAAACATG TCAGGTGGAT 180
GAGATCGTTA CACGATTACA TGA CTCCGTA TGT TTTCTAAA TCGCCGAGAG GAGCTTATTT 240
GANTTACAGA GATCTTGATT TGGGCTCGAC CAAAGGGATT AACACGGGTT TCGGAGATGC 300
AAGGAAATGG NNGGGTGAGN CTTTTTTTCAA AGGTAATTTT CAAGGGGTTA GGT TTTTGTT 360
AAAGGGGAGG TTTNNCCCAN CAAATTTTTTTT TTCAGGANCC GGCCANGNTT TTCCCCCCCC 420
TNTTTTTTNGG NCCCCAATCN AAANCCCCGT TT TAAAAGGG GGGCCATTTC NT TTTTNTNCA 480
NNTTAAAAGG G                                     491
```

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 407 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..407

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

```
ATTGTTCGT GAGGTTAACT TTGACTTTAG TCAACGGTAC GAAGCCTGGT GAGAATACGG 60
TTTTCGCGAC TTTCATTGGG ATGTATTTAG GCCGGTCGGA TAAGCTGTTG ACCGTNATGA 120
ACCGGGATTT CCCGGAGTTG AAGCTGAAGA AAACCGATTN TACCGAGATG AGATGGATCG 180
ATTCGGTTCT GTTTTGGGAC GATTATCCGG TTGGTACACC GACTTCTGTG CTACTAAATC 240
CGCTAGTCGC AAAAAAGTTG TTCATGAAAC GAAAATCGGA CTACGTGAAG CGTCTNATTT 300
TCGAGAACC CC GATCTCNNGT TTGATACTCA AGAAATTTGT AGAGGTTNNG AAAGTTAAAA 360
TNAATTTGGA TCCGCATTNN GGNANNNATG GTGAAACCCC NNGTTNT 407
```

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 360 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..360

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

ACGGCGTCGT ATTGGCCTAC AAAATAAACC TTGTTGAAGT CCCAGAAAAC GTCACCGTTT 60  
TCAGAATCTC CCGGACGTTA GAACAAAATG CGACGGATAT CATTACCGG TGGCAACAAG 120  
TTGCACCGAA GCTTCCCGAC GAGCTTTTCA TAAGANCAGT CATTGACGTA NAAACGGCAC 180  
TGTTTCATNN CTCAAAAGAC CGTCAGACAA CATTATAGC AATGTTTCTA GGAGACACGN 240  
CAACTCTACT GTCGATATTA AACCGGAGAT TCCCAGAATT GGGTTTGGTC CGGTCTGACT 300  
GTACCGNAAC AAGCNNTTGG ATCCAATCTG TGCTATTTTT GGGACAAATA TCCCAGGTTG 360

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 427 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..427

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

TCTTCACTGT CACCAAAACG TTAGAACAAG ACGCAAGATT GAAGACTATT TCTAAGTGGC 60  
AACAAATTTT ATCCAAGATT ATTGAAGAGA TACACATCCG AGTGGTACTC AGAGCAGCTG 120  
GAAATGATGG AAACAAGACT GTGACAATGA CCTACCTAGG TCAGTTTCTT GGCGAGAAAG 180  
GCACCTTGCT GAAGGTTATG GAGAAGGCTT TTCCAGAACT AGGGTTAACT CAAAAGGATT 240  
GTACTGAAAT GAGCTGGATT GAAGCCGCCC TTTTCCATGG TGGRTTCCA ACAGGKTCTC 300  
CTATTGAAAT TTTGCTTMAG CTCAAGTCGC CTYTAGGAAA AGRTTWCTTC AAAGCAACGK 360  
CGGATTTCGT TAAAGAACCT WTTCTGTGA TAGGGCTCAA AGGAATATTC AAAAGATTGA 420  
TTGAAGG 427

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 437 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana  
(B) STRAIN: ecotype Columbia

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 1..437

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

```
GTTGTACTAT CATNGAAGAT TAAGTTAGTC GATGTTCCGT CCACGGTCAC CGNGTTTAAA 60
GTCCAGAAAC ATNAGGAGAA AGAGGCCGTT AGGNTCATCA ACAAGTGGCA GTATGTTGCG 120
GATAAGGTCC CTGAAGATCT TTTCATCAGC GCAACGTTGG NGAGATCAAA CGGAAACTCT 180
GTGCAGGCTT TGTTTACTGG ACTCTATCTT GGNCCGGTGA ATAATNTCTT GGCCTTGATG 240
GAAGAAAAGT TTCCAGANTT AGGTCTTGAT ATCCAAGNCT GCACAGAGAT GAGTTGGGCT 300
GAATCTGCAC TCTGGTNTNC TGNTTTCNCT AAAGGAGAGN CTCCTTGGGT GTTCCNCGCG 360
GATCGGNAGC GGNCAATTTN TGGNCTTTCA AGGGGAAAGN CGGCTTTTNN CAAGAACCCG 420
NTACCCGGGG TTCAATT 437
```

(2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 441 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 1..441

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

```
GCGGACCCCTA TAGATCANNA TGTGCTACTG ANAGAAGAGG AAGCCAAGAA CAAGCCGGAG 60
ACAGATAAAT ATCTGAAATG GGNCGATANC GTTTACGAAT TTATGACNCC ATATGTTTCG 120
```

AAATCTCCAA GAGGAGCTTA TGTCAATTTT AAGGATATGG ATTTGGGTAT GTATCTTGGA 180  
AAGAAGAAGA CAAAGTACGA GGAAGGAAAG AGTTGGGGAG TGAAGTATTT CAAGAACAAT 240  
TTCGAGAGAT TGGTGAGAGT GAAGACTAGG GTTGATCCAA CAGATTTCTT CTGCGATGAA 300  
CAGAGCATTG CTCTGGTGAA CAAAGTTACC TGAAGATATC ATTTGAAGTT TTTTATTAGT 360  
CCCTTTTCTC TGTGAAATCA TCTGTGCGTG TTGAATATTA TGCCTCAAGT GTGTAACTTA 420  
TGTGTGTGAT TGTGAATTGT G 441

(2) INFORMATION FOR SEQ ID NO: 33:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 502 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Arabidopsis thaliana
  - (B) STRAIN: ecotype Columbia
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 2..502
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

CTGGCTTAAC ACAACGTCGT TTTGGGCCAA TTACCCGGCG GGTACACCCA AGAGCATCCT 60  
TCTAGATAGG CCTCCGACGA ATTCAGTGTC ATTTAAGAGT AAATCGGATT TTGTCAAAAA 120  
ACCAATACCC AAAAAAGGTT TAGAGAAGCT TTGGAAGACA ATGTTTAAAT TCAACAGTAG 180  
CGTCTCGTTG CAATTCAACC CTTACGGTGG AGTGATGGAC CGGATTCCGG CAACGGCCAC 240  
CGCTTTTCCT CATCGGAAAG GAACTTGTT CAAGGTTCAA TACNCTACGA TGTGGTTTGA 300  
CGCAAACGCC ACACAGAGTA GCCNGGCTAT GATGAATGAG CTTTTTGAGG TGGCGGGACC 360  
GTACGTGNGT CAAGTAAACC CGAGANANGG CTTCTTTTAA NTTACAGAGNC CATCGNTNTT 420  
NGGAGCAANN CCAAGTGGGG GGGNCCAACC GGGGGNTNAA ANCNNAGNTC TTNGGGGGCC 480  
CAGAATTTC TTNGGGGAAT TT 502

(2) INFORMATION FOR SEQ ID NO: 34:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 400 base pairs
  - (B) TYPE: nucleic acid

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana  
(B) STRAIN: ecotype Columbia

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 2..400

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

```
NGGGAATTGC NCGAGGNAAG TTGTACCCAA TTCCTGGACC ACCATTGGTT TCCCAAGAAN 60
CCCGAGACAA CCGTTTTTCA ATNACCGTGA TGTTGATTG GGTATTAATT CTCATAATGG 120
TAAAATCAGT AGTTATGTGG AAGGTAAACG TTACGGGAAG AAGTATTTTCG CAGGTAATTT 180
CGAGAGATTG GTGAAGATTA AGACGAGAGT TGATAGTGGT AATTTCCTTA GGAACGAACA 240
GAGTATTCCT GTGTTACCAT AAGTGTATTT ATTTGATTAT TGGTTAGTGA AATTTGTTGT 300
TGTATAATGA TTATATGTCG TATTTTTATT TATTATTAGT AATTTATAAA GTTTGATATT 360
AAATACAAAT AGTATAATAA GATAGTTTCT TTTAGTAAAA 400
```

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 383 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana  
(B) STRAIN: ecotype Columbia

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 2..383

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

```
CAACTCTAAT GGAACACCT ACTTCGATCG AATGTCGATG GGGGAAGAGC TTTTCTGGGC 60
GGTTCGAGGA GGTGGAGCCG CGAGTTTCGG CATCGTGATG GGATACAAA TCCGGTTGGT 120
```

TCCGGTTCCG GAGAAAGTTA CGGTTTTTAG CGTCGGAAAA ACCGTCGGAG AAGGAGCCGT 180  
 TGATCTTATA ATGAAGTGGC AGAACTTCTC TCATAGTACG GNTCGGAATT TTTTTGTGAA 240  
 GCTGANTTTT GANTTTAGTC AACGGTGCAA AGCCGGGTGA AAAAAAGGTT TTAGNGNCTT 300  
 TCANTTTGGN TGNAANCTTG GGGGTTTTAT NAGAACGGTT AACCAGGATT NANCCCGNGT 360  
 TTTCCCGGGG TTAAACCTT NGG 383

(2) INFORMATION FOR SEQ ID NO: 36:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 354 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Arabidopsis thaliana
  - (B) STRAIN: ecotype Columbia
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 1..354
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

ATCAATGTTC TTAATAACG TACACGAGCA TCGTTGGCTT TCAAGGCTAA ATCTGATTTT 60  
 NTTCAAGAAC CGATNCCTAA AACCGCGATT TCGAAGCTTT GGAGACGGTT GCAAGAACCG 120  
 GAAGCAGAGC ATGCTCAGCT AATTNCACN CCATTGCTG GTAAATGAG TNAGATTGCA 180  
 GATTACGAAA CACCATTTC GCATAGGAAG GGAATATAT ATNAGATTCA GTACTTGAAT 240  
 TACTGGAGAG GAGACGTGAA AGAGAAGTAT ATTGAGATNG GTGGAGGAGA GTTTACGGTT 300  
 GNTATNAGTA AGTTTTTTGG CGAAGTNTNC CNAGAGGNGN CTTNNTNTAA ACCT 354

(2) INFORMATION FOR SEQ ID NO: 37:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 403 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO



(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..403

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

```
TTTTTTAGTA CACTAATAAT CAAATGGAAT GAGAAATGAA GCCACAAAAG TATCTGCAAT 60
CAAAATATCC TGCTATCTCC ATCTCAAGCT CTCAATAGTA TCCTCTCCGA AAGTGAAATC 120
AACATTTCAA ACTCTATTTC TTGGTGAAT CGATAGACTG ATTCCTCTGA TGAACCAGAA 180
GTTTCCGGAA CTCGGCTTAC GATCTCAAGA CTGTTCGGAA ATGAGCTGGA TCGAATCGAT 240
AATGTTCTTC AACTGGAGAT CAGGACAGCC GTTAGAGATT TTGCTCAACA GAGACCTAAG 300
GATTGAGGA TCAGTATTTC AAAGCAAAGT CAGGATTATG GTTCAAAAAC CCGTTCCTGA 360
AAACGTTTTT CGAAGAGGTA TCCAAGGGGT TTCTCGAGCA AGT 403
```

(2) INFORMATION FOR SEQ ID NO: 38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 260 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..260

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

```
GAGATGAGTT GGATTAANTC TGTACTCTGG TTTGCTGATT TCCCTAAAGG AGAATCTCTT 60
NGTGTTCTCA CGAATCGTAA GCGTACATCT CTATCTTTNA AAGGCAAAGA TGATTTTATC 120
CAAGAACCGA TACCCGAGGC TGCAATTNAA GAGATATGGA GGCGATTAGA AGCCCCCNAG 180
GCTCGGCTTG GAAAGATCAT ATTA ACTCCA TTTGGTGGGA AAATNAGTGA AATGGCAGAG 240
TACGTANCAC CATTCCCACA 260
```

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 605 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..605

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

```
CTCTTGCATA TTCGCTGCAA GGATGGGAAA TTCAAAACCA CTCCCTACAA TTTTTTGTAT 60
TATAGTTTCA GTCTTGATT TTTAATTCTA TTGCATAACA CCAACTTCTT CATCAGCCTC 120
CATCCAAGAT CAATTCATAA ACTGTGTCAA AAGAAACACA CATGTTTCTT TTCCACTCGA 180
GAAAACGTTA TTCACCCCTG CGAAAAACGT CTCTTTGTTC AACCAAGTCC TTGANTCGAC 240
GGCTCAAAAT CTCCAGTTCT TGGCAAAATC CATGCCTAAA CCGGGRTTCA TATTCAGACC 300
GATTCACCAG TCTCAAGTCC AAGSTTCCAT CATTTGTTCA AMGRAACTCG GGNITCATT 360
TNGTGTTTGA NGTGGCGGTC ACGATTTTCG AGGCCTTTGT NTTTATGTTT CACGGTTTGA 420
AAAAACCGTT TATATTACTC GGCCTGTCAA ANTTGNANNC AAAATCANAT GTTGGATATT 480
GNATTCCAAA TAGGTNCTTG GGGTNAACCT GGTGGCTANC GTTTGGTGAG CTTTTACTTT 540
CAAGAATTTG CANGNGGANG TGCAAAGATT CCATGGGATT TCCCGGGGGG TTTNTTGCAC 600
AATGT 605
```

(2) INFORMATION FOR SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 464 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana

(B) STRAIN: ecotype Columbia

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 2..464

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

```
AAACAAAAAC TCTTCCATTT GGCTTCTCTC TTGCATATTC GTTGCAAGGA TGGGAAATTC 60
AAAACCACTC CCTACAATTN CTTGTATTAT CGTTTCAGTC TTGTATTTTN NATTCTATTG 120
CATAACACCA ACTTCTTCAT CAGCCTCCAT CCAAGNTCAA TTCATAAACT GTGTCAAAAAG 180
GAACACACAT GTTTCTTTTC CACTCGAGNA AACGGTATTC ACTCCTGCGG AAAACGGCTC 240
TNTTATTCAA CGGGTCCNTG AATCGACGGG TCAAATCTC CAGTTCTTGG NAAAATCCAT 300
GNCTAAACCG GGGTTCATAT TCAGGCCGGT TCACCAGTCT CAAGTCCAAG NTTCCATCAT 360
TTGTTCAAAG GAACTCGGGA TTCATTTCCG CGNTAGAAGT GGCGGGCANN GGTTTCGGGG 420
CCTGTCTNTT GNTTANGGGN AGGAAAACCG GTTNTATTNC TCGG 464
```

(2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 386 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana

(B) STRAIN: ecotype Columbia

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..386

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

```
TCGGGAGCCC ANGNTAAATT ANNTGAAAAT GGGGNCGNAT ANCCGTTTAC NGAATTTTAT 60
GACNCCCAAT ATGTTTCGAA ATCTCAAAGA NNGGGANCTT ATGTCAATTT CAAGGATATG 120
GATTTGGGTA TGTATCTTGG AAAGNAGAAG ACAAAGTACG AGGAAGGAAA GAGTTGGGGA 180
GTGAAGTATT TCAAGAACAA TTTCGAGAGA TTGGTGAGAG TGAAGACTAG GGTGATCCN 240
ACAGATTTCN TCTGCGATGA ACAGAGCATT CCTCTGGTGN ACAAAGTTAC CTGAAGATAT 300
CATTTGAAGT TTTTATTAG TCCCTTTTCT CTGTGAAATC ATCTGTGCGT GTTGAATANT 360
```

## (2) INFORMATION FOR SEQ ID NO: 42:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 377 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..377

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

TACCATAGGG AGGTGGTGNA AGATTTTGTGTA TGTAGNCTTA GGGGAAGGCG AGTAGTATGG 60  
TGGTGGTGGG GAGCTGTAAA CGTATGGTGG TGGTGGAGAT TTGTATGTGG GCTGGTTAAC 120  
TTCATTGAAG CTAAAATCTG GGGACCTAAG TACTTCAAAG GCAATTTTGA CAGATTGGTG 180  
AAGATTAAAA CCAAGGTTGA TCCAGAGAAC TTCTTCAGGC ACGAGCAGAG TATCCCACCT 240  
ATGCCCTACT AGAAGCTAGG TTCATGAAAC CAATAACATT ATCAAAAATA AGAATAAATG 300  
ATAATTGTAT ACAACATGAT TCGTCTTTCT TTATTCAGA CAATGTGGAC ACTACTCTAA 360  
ATAAAATGTC ATTTACC 377

## (2) INFORMATION FOR SEQ ID NO: 43:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 377 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

## (ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 1..377

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

```
TACCATAGGG AGGTGGTGNA AGATTTTGTA TGTAGNCTTA GGGGAAGGCG AGTAGTATGG 60
TGGTGGTGGG GAGCTGTAAA CGTATGGTGG TGGTGGAGAT TTGTATGTGG GCTGGTTAAC 120
TTCATTGAAG CTAAAATCTG GGGACCTAAG TACTTCAAAG GCAATTTTGA CAGATTGGTG 180
AAGATTAAAA CCAAGGTTGA TCCAGAGAAC TTCTTCAGGC ACGAGCAGAG TATCCACCT 240
ATGCCCTACT AGAAGCTAGG TTCATGAAAC CAATAACATT ATCAAAAATA AGAATAAATG 300
ATAATTGTAT ACAACATGAT TCGTCTTTCT TTATTCAGA CAATGTGGAC ACTACTCTAA 360
ATAAAATGTC ATTTACC 377
```

(2) INFORMATION FOR SEQ ID NO: 44:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 346 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana  
(B) STRAIN: ecotype Columbia

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 2..346

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

```
GAGCTGTGGA TATGGTCACA AATGGCAATC GGTGGTCCG AAAACTGATC CGAATCTTTT 60
TATGAGAATN TTGATTCAAC CAGTGACGAG GAAGAAGGTA AAGACTGTGA GAGCTTCTNT 120
GGTTGCCCTN TTTTNAGGCN AGACAGATGA AGTTTTTGCT TTCCTTAGTA AGGAGTTTCC 180
TGAATTGGGT TTAAAGAAGG AGAATTNTTC GGAGATGACT TGGTTTCANT CTGCTTTATG 240
GTGGGACAAT CGTCTTAATG CTAATCAGGT TGATCCTAAA GTNTTCTTG ATCGGAATCT 300
CGATACCTCG AGTTTCGGTA AGAGGAAATC GGATTACGTC GCGACT 346
```

(2) INFORMATION FOR SEQ ID NO: 45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 261 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..261

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

```

ATGGGGTGAG ACTTATTTCA AAGGTAATTT CAAGAGATTA GGTTCGGTA AAGGGAAGNT   60
TGATCCAACA AATTTCTTCA GGAACGAACA GAGTATTCCT CCTCTGTTT GAGTCCTCAA  120
TACAAAACCA GATATAAAAG ATGTCATTTC ATTTTTTCAA TTATAATAGA TAATGTAAC  180
TTCTGCTACA ATTGTAAAAG TGAGATGTAC CCAATACGGT TTAAGCGGAC CGAGAATAGT  240
CAATTCAAAG ACCAAATTCT G                                           261

```

(2) INFORMATION FOR SEQ ID NO: 46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 478 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..478

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

```

GCTCAAAGGA CTAACCATGA AACTTCCTC AAGTGTCTCT CTCACCGANT CAACGAGGAC   60
GACTCAAGAN TTATACACAC ATCAAAAGAT CCTTCGTATT TNTCAATCTT GATTTCTTCC  120
ATACAAAATC CAAGTTTCTC TGTTCGAA ACACCTAAAC CGGTTTCAAT CATCACTCCG  180

```

GTTCAAGCCA CCGATGTTCA ATCTACGNTT AAATNCGCAC GGNCTTCACG GGTATACACA 240  
 ATCAGGGCTA GGAGTGGTNG TCATGACTAC GGAGGTTTAT CTTTACATTG GCTTAAAAAN 300  
 CANNCCGTTT GTTNNTCATT GATTTNNAGA AATCTTCCGG GCTTATTAA CATNTAAGAT 360  
 GTTTGATAAN CCGGNNCCNG TTTGGGGTTC AAATCCCGGT GGCTTACAAA NTTNGGGGGA 420  
 ATGTGNCCTA TGAGGTTTGG AAAATTAANG CAAAATNTTT TGGGCCTTCC CGGCCGGT 478

(2) INFORMATION FOR SEQ ID NO: 47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 579 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..579

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

GGCCGTTAGG ATCATCAAGA AATGGCAATA TGCTGCAGAT AAGGTTCTTG ATGATCTTTT 60  
 CATTAGGACA ACATTGGAGA GATCAAACAA GAACGCAGTA CACGCTTTGT TCACTGGACT 120  
 ATATATTGGT CCGGTGAACA ATCTATTGGC GTTGATGGAA GAAAAGTTTC CGGAACTAGG 180  
 TCTTGAGAAA GAAGGTTGTG AAGAGATGAG TTGGATTGAG TCTGTACTCT GGTTCGCTGA 240  
 TTTCCCTAAA GGAGAATCTC TTGGTGTCT CACGAATCGT GAGCGTACAT CTCTATCTTT 300  
 CAAAGGCAAA GATGATTTTG TCCAAGAACC GATACCCGAG GCTGCAATTC AAGAGATATG 360  
 GAGGCGATTA GAAGCCCCCG AGGCTCGGCT TGGAAGATC ATATTAAGTC CATTTGGGTG 420  
 NGGNAAAATG AGTGAAATGG CAGAGNCCGA ACCACCAATT CCCACANNCG AGGGAGGGGA 480  
 ACCCCTNTGN GGNTCAGAAAT GTGGTTCCTG GNNNNNAAGN GGGNGCCAGN ACCAANCCGG 540  
 GNCNGTAAAN CNTGNAATGG GCCNAACCCG TNCCGGATT 579

(2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 252 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Oryza sativa  
(B) STRAIN: Nipponbare, subsp. japonica  
(D) DEVELOPMENTAL STAGE: etiolated shoot (8 days old)

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 3..252

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

```
TGTCCTGGAA GGTCCGCCTC GTGCAGGTTN CGACGACGGT GACGGTGTTC GTCGTCGGGA 60
GGAACGTCTGA CCAGGGCGCC GCNGACGTCG TCGCCAGATG GCAAGACGTC GCGCCGAGCC 120
TCCCTCCCGA GCTCACCATA CGGGTGATCG TNCGAGGGCA GCGCGCCACG TTCCAGTCGC 180
TGTACCTCGG CTCGTGCGCC GACCTGGTGC CGACGATGAG CAGCATGTTC CCGGAGCTCG 240
GGATGACGAT TG 252
```

(2) INFORMATION FOR SEQ ID NO: 49:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Lactuca sativa  
(B) STRAIN: lollo bionda

(ix) FEATURE:

(A) NAME/KEY: Modified-site  
(B) LOCATION: 12  
(D) OTHER INFORMATION: /label= Ambiguous  
/note= "Xaa = Cys or Ser"

(ix) FEATURE:

(A) NAME/KEY: Modified-site  
(B) LOCATION: 20..21  
(D) OTHER INFORMATION: /label= ambiguous  
/note= "Xaa-Xaa probably is Ser-Phe"



[illegible]

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(A) ORGANISM: Lactuca sativa  
(B) STRAIN: lollo bionda

(A) NAME/KEY: Modified-site  
(B) LOCATION: 1  
(D) OTHER INFORMATION: /label= ambiguous  
/note= "Xaa = probably Ser"

(A) NAME/KEY: Modified-site  
(B) LOCATION: 3  
(D) OTHER INFORMATION: /label= unknown

(A) NAME/KEY: Modified-site  
(B) LOCATION: 5  
(D) OTHER INFORMATION: /label= ambiguous  
/note= "Xaa = probably Ser"

(A) NAME/KEY: Modified-site  
(B) LOCATION: 12  
(D) OTHER INFORMATION: /label= ambiguous  
/note= "Xaa = probably Trp"

(A) NAME/KEY: Modified-site  
(B) LOCATION: 24  
(D) OTHER INFORMATION: /label= ambiguous  
/note= "Xaa = probably Tyr"

Xaa Ile Xaa Val Xaa Ile Glu Asp Glu Thr Ala Xaa Val Gln Ala Gly  
1 5 10 15

Ala Thr Leu Gly Glu Val Tyr Xaa  
20

(2) INFORMATION FOR SEQ ID NO: 51:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 14 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Lactuca sativa
  - (B) STRAIN: lollo bionda
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:  

Ala	Asp	Pro	Ser	Phe	Pro	Leu	Ser	Gly	Gln	Leu	Tyr	Tyr	Pro
1				5					10				

(2) INFORMATION FOR SEQ ID NO: 52:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 32 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

ACTTCTACTT CTATTATTGA TAGGTTTACT CA

32

(2) INFORMATION FOR SEQ ID NO: 53:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 405 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Lactuca sativa
  - (B) STRAIN: lollo bionda
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 1..405

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

ACT TCT ACT TCT ATT ATT GAT AGG TTT ACT CAA TGT CTA AAC AAC CGA	48
Thr Ser Thr Ser Ile Ile Asp Arg Phe Thr Gln Cys Leu Asn Asn Arg	
1 5 10 15	
GCT GAC CCT TCT TTC CCG CTC AGT GGA CAA CTT TAC ACT CCC GAT AAC	96
Ala Asp Pro Ser Phe Pro Leu Ser Gly Gln Leu Tyr Thr Pro Asp Asn	
20 25 30	
TCC TCT TTT CCA TCC GTC TTG CAA GCT TAC ATC CGG AAC CTC CGA TTC	144
Ser Ser Phe Pro Ser Val Leu Gln Ala Tyr Ile Arg Asn Leu Arg Phe	
35 40 45	
AAT GAA TCC ACG ACT CCC AAA CCC ATC TTA ATC ATC ACC GCC TTA CAC	192
Asn Glu Ser Thr Thr Pro Lys Pro Ile Leu Ile Ile Thr Ala Leu His	
50 55 60	
CCT TCA CAC ATT CAA GCA GCT GTT GTG TGC GCC AAA ACA CAC CGC CTG	240
Pro Ser His Ile Gln Ala Val Val Cys Ala Lys Thr His Arg Leu	
65 70 75 80	
CTA ATG AAA ACC AGA AGC GGA GGC CAT GAT TAT GAG GGG CTT TCC TAT	288
Leu Met Lys Thr Arg Ser Gly Gly His Asp Tyr Glu Gly Leu Ser Tyr	
85 90 95	
GTG ACC AAT TCG AAC CAA CCC TTT TTT GTT GTT GAC ATG TTC AAC TTA	336
Val Thr Asn Ser Asn Gln Pro Phe Phe Val Val Asp Met Phe Asn Leu	
100 105 110	
CGC TCC ATA AAC GTG AGT ATT GAA GAT GAA ACT GCA TGG GTC CAA GCC	384
Arg Ser Ile Asn Val Ser Ile Glu Asp Glu Thr Ala Trp Val Gln Ala	
115 120 125	
GGC GCC ACC CTC GGA GAA GTT	405
Gly Ala Thr Leu Gly Glu Val	
130 135	

(2) INFORMATION FOR SEQ ID NO: 54:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 135 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

Thr Ser Thr Ser Ile Ile Asp Arg Phe Thr Gln Cys Leu Asn Asn Arg	
1 5 10 15	
Ala Asp Pro Ser Phe Pro Leu Ser Gly Gln Leu Tyr Thr Pro Asp Asn	
20 25 30	
Ser Ser Phe Pro Ser Val Leu Gln Ala Tyr Ile Arg Asn Leu Arg Phe	
35 40 45	
Asn Glu Ser Thr Thr Pro Lys Pro Ile Leu Ile Ile Thr Ala Leu His	
50 55 60	

Pro Ser His Ile Gln Ala Ala Val Val Cys Ala Lys Thr His Arg Leu  
65 70 75 80

Leu Met Lys Thr Arg Ser Gly Gly His Asp Tyr Glu Gly Leu Ser Tyr  
85 90 95

Val Thr Asn Ser Asn Gln Pro Phe Phe Val Val Asp Met Phe Asn Leu  
100 105 110

Arg Ser Ile Asn Val Ser Ile Glu Asp Glu Thr Ala Trp Val Gln Ala  
115 120 125

Gly Ala Thr Leu Gly Glu Val  
130 135

(2) INFORMATION FOR SEQ ID NO: 55:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

CACGTTTATG GAGCGTAAGT TGAAC

25

(2) INFORMATION FOR SEQ ID NO: 56:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

CACCCTTCAC ACATTCAAGC AGC

23

(2) INFORMATION FOR SEQ ID NO: 57:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1981 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Lactuca sativa
- (B) STRAIN: lollo bionda

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 7..1626

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 372
- (D) OTHER INFORMATION: location 372 may be "C" or "G"

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 379
- (D) OTHER INFORMATION: location 379 may be "A" or "G"

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 786
- (D) OTHER INFORMATION: location 786 may be "C" or "T"

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 1105...1106
- (D) OTHER INFORMATION: location 1105...1106 may be "AG",  
"GA", "GG" or "AA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

ACAAAA ATG GCA ATT ACC TAT TCT TTC AAC TTC AAA TCT TAT ATT TTT	48
Met Ala Ile Thr Tyr Ser Phe Asn Phe Lys Ser Tyr Ile Phe	
1 5 10	
CCT CTC CTC CTT GTC TTG CTC TCT ACC CAT TCA TCA GCG ACT TCA ACT	96
Pro Leu Leu Leu Val Leu Leu Ser Thr His Ser Ser Ala Thr Ser Thr	
15 20 25 30	
TCC ATT ATA GAT CGC TTC ACC CAA TGT CTA AAC AAC CGA GCT GAC CCT	144
Ser Ile Ile Asp Arg Phe Thr Gln Cys Leu Asn Asn Arg Ala Asp Pro	
35 40 45	
TCT TTC CCG CTC AGT GGA CAA CTT TAC ACT CCC GAT AAC TCC TCT TTT	192
Ser Phe Pro Leu Ser Gly Gln Leu Tyr Thr Pro Asp Asn Ser Ser Phe	
50 55 60	
CCA TCC GTC TTG CAA GCT TAC ATC CGG AAC CTC CGA TTC AAT GAA TCC	240
Pro Ser Val Leu Gln Ala Tyr Ile Arg Asn Leu Arg Phe Asn Glu Ser	
65 70 75	
ACG ACT CCC AAA CCC ATC TTA ATC ATC ACC GCC TTA CAC CCT TCA CAC	288
Thr Thr Pro Lys Pro Ile Leu Ile Ile Thr Ala Leu His Pro Ser His	
80 85 90	
ATT CAA GCA GCT GTT GTG TGC GCC AAA ACA CAC CGC CTG CTA ATG AAA	336
Ile Gln Ala Ala Val Val Cys Ala Lys Thr His Arg Leu Leu Met Lys	
95 100 105 110	

ACC AGA AGC GGA GGC CAT GAT TAT GAG GGG CTT TCS TAT GTG RCC AAT	384
Thr Arg Ser Gly Gly His Asp Tyr Glu Gly Leu Ser Tyr Val Xaa Asn	
115 120 125	
TCG AAC CAA CCC TTT TTT GTT GTT GAC ATG TTC AAC TTA CGC TCC ATA	432
Ser Asn Gln Pro Phe Phe Val Val Asp Met Phe Asn Leu Arg Ser Ile	
130 135 140	
AAC GTG AGT ATT GAA GAT GAA ACT GCA TGG GTC CAA GCT GGT GCG ACT	480
Asn Val Ser Ile Glu Asp Glu Thr Ala Trp Val Gln Ala Gly Ala Thr	
145 150 155	
CTT GGT GAA GTC TAC TAC CGA ATA GCA GAG AAA AGC AAC AGT CAT GCT	528
Leu Gly Glu Val Tyr Tyr Arg Ile Ala Glu Lys Ser Asn Ser His Ala	
160 165 170	
TTT CCG GCT GGC GTT TGC CCT ACT GTT GGA GTT GGT GGC CAT TTT AGT	576
Phe Pro Ala Gly Val Cys Pro Thr Val Gly Val Gly Gly His Phe Ser	
175 180 185 190	
GGT GGT GGT TAT GGT AAC TTG ATG GGA AAA TAC GGC CTT TCT GTT GAC	624
Gly Gly Gly Tyr Gly Asn Leu Met Gly Lys Tyr Gly Leu Ser Val Asp	
195 200 205	
AAT ATT GTC GAT GCT CAG TTA ATC GAT GTG AAT GGT AAA CTT CTG AAT	672
Asn Ile Val Asp Ala Gln Leu Ile Asp Val Asn Gly Lys Leu Leu Asn	
210 215 220	
CGG AAA TCA ATG GGT GAA GAT CTT TTT TGG GCC ATC ACA GGT GGT GGT	720
Arg Lys Ser Met Gly Glu Asp Leu Phe Trp Ala Ile Thr Gly Gly Gly	
225 230 235	
GGT GTC AGC TTT GGT GTG GTT GTA GCG TAC AAG ATC AAA CTG GTT CGT	768
Gly Val Ser Phe Gly Val Val Val Ala Tyr Lys Ile Lys Leu Val Arg	
240 245 250	
GTT CCT ACC ACT GTG ACY GTT TTT AAC GTA CAA AGA ACA TCC GAG CAG	816
Val Pro Thr Thr Val Val Phe Asn Val Gln Arg Thr Ser Glu Gln	
255 260 265 270	
AAC CTA AGC ACC ATA GCC CAC CGA TGG ATA CAA GTT GCG GAT AAG CTC	864
Asn Leu Ser Thr Ile Ala His Arg Trp Ile Gln Val Ala Asp Lys Leu	
275 280 285	
GAT AAT GAC CTT TTC CTT CGA ATG ACC TTT AAC GTG ATA AAC AAC ACA	912
Asp Asn Asp Leu Phe Leu Arg Met Thr Phe Asn Val Ile Asn Asn Thr	
290 295 300	
AAT GGC GAA AAG ACG ATA CGT GGT TTG TTT CCA ACA CTG TAC CTC GGA	960
Asn Gly Glu Lys Thr Ile Arg Gly Leu Phe Pro Thr Leu Tyr Leu Gly	
305 310 315	
AAC TCT ACC GCT CTT GTT GCC CTC CTG AAC AAG GAT TTC CCT GAA TTA	1008
Asn Ser Thr Ala Leu Val Ala Leu Leu Asn Lys Asp Phe Pro Glu Leu	
320 325 330	
GGT GTA GAA ATT TCA GAT TGT ATT GAA ATG AGT TGG ATC GAG TCT GTT	1056
Gly Val Glu Ile Ser Asp Cys Ile Glu Met Ser Trp Ile Glu Ser Val	
335 340 345 350	

CTT TTC TAC ACA AAC TTC CCC ATT GGT ACT CCG ACC ACT GCT CTT CTA 1104  
 Leu Phe Tyr Thr Asn Phe Pro Ile Gly Thr Pro Thr Thr Ala Leu Leu  
 355 360 365

RRC CGT ACA CCT CAA AGA CTA AAC CCA TTC AAA ATC AAA TCT GAT TAC 1152  
 Xaa Arg Thr Pro Gln Arg Leu Asn Pro Phe Lys Ile Lys Ser Asp Tyr  
 370 375 380

GTA AAA AAC ACT ATT TCC AAA CAG GGA TTC GAA TCC ATA TTT GAA AGG 1200  
 Val Lys Asn Thr Ile Ser Lys Gln Gly Phe Glu Ser Ile Phe Glu Arg  
 385 390 395

ATG AAA GAA CTC GAA AAC CAA ATG CTA GCT TTC AAC CCT TAT GGT GGA 1248  
 Met Lys Glu Leu Glu Asn Gln Met Leu Ala Phe Asn Pro Tyr Gly Gly  
 400 405 410

AGA ATG AGC GAA ATT TCC GAA TTT GCA AAG CCT TTT CCC CAT CGA TCA 1296  
 Arg Met Ser Glu Ile Ser Glu Phe Ala Lys Pro Phe Pro His Arg Ser  
 415 420 425 430

GGG AAT ATA GCG AAG ATC CAA TAC GAA GTA AAC TGG GAT GAA CTT GGC 1344  
 Gly Asn Ile Ala Lys Ile Gln Tyr Glu Val Asn Trp Asp Glu Leu Gly  
 435 440 445

GTT GAA GCA GCC AAT CGG TAC TTG AAC TTC ACA AGG GTG ATG TAT GAT 1392  
 Val Glu Ala Ala Asn Arg Tyr Leu Asn Phe Thr Arg Val Met Tyr Asp  
 450 455 460

TAT ATG ACT CCG TTT GTT TCT AAG AAC CCC AGG GAA GCA TTT CTG AAC 1440  
 Tyr Met Thr Pro Phe Val Ser Lys Asn Pro Arg Glu Ala Phe Leu Asn  
 465 470 475

TAC AGG GAT TTA GAT ATT GGT GTC AAC AGT CAT GGC AAG AAT GCT TAC 1488  
 Tyr Arg Asp Leu Asp Ile Gly Val Asn Ser His Gly Lys Asn Ala Tyr  
 480 485 490

GGT GAA GGA ATG GTT TAT GGG CAC AAG TAT TTC AAA GAG ACG AAT TAT 1536  
 Gly Glu Gly Met Val Tyr Gly His Lys Tyr Phe Lys Glu Thr Asn Tyr  
 495 500 505 510

AAG AGG CTA ACG ATG GTG AAG ACG AGG GTT GAT CCT AGC AAT TTT TTT 1584  
 Lys Arg Leu Thr Met Val Lys Thr Arg Val Asp Pro Ser Asn Phe Phe  
 515 520 525

AGG AAT GAG CAA AGT ATC CCA ACT TTG TCA TCT TCA TGG AAG 1626  
 Arg Asn Glu Gln Ser Ile Pro Thr Leu Ser Ser Ser Trp Lys  
 530 535 540

TAAATTCTAA ATTCACCTTGT GAAATTGAAT AAAAGTATGG CTTTTTCAAG GTCATGGTAT 1686

CCAGATTCAG ATGATATTGA TATAATTTTG ACTTGTTATTT ATACAAACAA AATTATATTA 1746

TATTTTTCTG AATTTAGATT TTCCATTCTT TGGAAAAATA TACGAACATT GATGTTGATA 1806

TTTTAAAGAA TTATAGATTT TGAACATTGT GAACAATGAA TAAACCGAGG ACTTCCCTTG 1866

GGTTTTTTTTT ATAAGTATGT AATAGCATGT CTTTAATCAA GATAACCGAT CATTGGATGC 1926

AATTTATTAT TATAACCTT ATTTAAAAA AAAAAAAAAA AAAAAAAAAA AAAAA 1981

(2) INFORMATION FOR SEQ ID NO: 58:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 540 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

Met Ala Ile Thr Tyr Ser Phe Asn Phe Lys Ser Tyr Ile Phe Pro Leu  
1 5 10 15  
Leu Leu Val Leu Leu Ser Thr His Ser Ser Ala Thr Ser Thr Ser Ile  
20 25 30  
Ile Asp Arg Phe Thr Gln Cys Leu Asn Asn Arg Ala Asp Pro Ser Phe  
35 40 45  
Pro Leu Ser Gly Gln Leu Tyr Thr Pro Asp Asn Ser Ser Phe Pro Ser  
50 55 60  
Val Leu Gln Ala Tyr Ile Arg Asn Leu Arg Phe Asn Glu Ser Thr Thr  
65 70 75 80  
Pro Lys Pro Ile Leu Ile Ile Thr Ala Leu His Pro Ser His Ile Gln  
85 90 95  
Ala Ala Val Val Cys Ala Lys Thr His Arg Leu Leu Met Lys Thr Arg  
100 105 110  
Ser Gly Gly His Asp Tyr Glu Gly Leu Ser Tyr Val Thr Asn Ser Asn  
115 120 125  
Gln Pro Phe Phe Val Val Asp Met Phe Asn Leu Arg Ser Ile Asn Val  
130 135 140  
Ser Ile Glu Asp Glu Thr Ala Trp Val Gln Ala Gly Ala Thr Leu Gly  
145 150 155 160  
Glu Val Tyr Tyr Arg Ile Ala Glu Lys Ser Asn Ser His Ala Phe Pro  
165 170 175  
Ala Gly Val Cys Pro Thr Val Gly Val Gly Gly His Phe Ser Gly Gly  
180 185 190  
Gly Tyr Gly Asn Leu Met Gly Lys Tyr Gly Leu Ser Val Asp Asn Ile  
195 200 205  
Val Asp Ala Gln Leu Ile Asp Val Asn Gly Lys Leu Leu Asn Arg Lys  
210 215 220  
Ser Met Gly Glu Asp Leu Phe Trp Ala Ile Thr Gly Gly Gly Gly Val  
225 230 235 240  
Ser Phe Gly Val Val Val Ala Tyr Lys Ile Lys Leu Val Arg Val Pro  
245 250 255  
Thr Thr Val Thr Val Phe Asn Val Gln Arg Thr Ser Glu Gln Asn Leu  
260 265 270



Ser Thr Ile Ala His Arg Trp Ile Gln Val Ala Asp Lys Leu Asp Asn  
 275 280 285  
 Asp Leu Phe Leu Arg Met Thr Phe Asn Val Ile Asn Asn Thr Asn Gly  
 290 295 300  
 Glu Lys Thr Ile Arg Gly Leu Phe Pro Thr Leu Tyr Leu Gly Asn Ser  
 305 310 315 320  
 Thr Ala Leu Val Ala Leu Leu Asn Lys Asp Phe Pro Glu Leu Gly Val  
 325 330 335  
 Glu Ile Ser Asp Cys Ile Glu Met Ser Trp Ile Glu Ser Val Leu Phe  
 340 345 350  
 Tyr Thr Asn Phe Pro Ile Gly Thr Pro Thr Thr Ala Leu Leu Ser Arg  
 355 360 365  
 Thr Pro Gln Arg Leu Asn Pro Phe Lys Ile Lys Ser Asp Tyr Val Lys  
 370 375 380  
 Asn Thr Ile Ser Lys Gln Gly Phe Glu Ser Ile Phe Glu Arg Met Lys  
 385 390 395 400  
 Glu Leu Glu Asn Gln Met Leu Ala Phe Asn Pro Tyr Gly Gly Arg Met  
 405 410 415  
 Ser Glu Ile Ser Glu Phe Ala Lys Pro Phe Pro His Arg Ser Gly Asn  
 420 425 430  
 Ile Ala Lys Ile Gln Tyr Glu Val Asn Trp Asp Glu Leu Gly Val Glu  
 435 440 445  
 Ala Ala Asn Arg Tyr Leu Asn Phe Thr Arg Val Met Tyr Asp Tyr Met  
 450 455 460  
 Thr Pro Phe Val Ser Lys Asn Pro Arg Glu Ala Phe Leu Asn Tyr Arg  
 465 470 475 480  
 Asp Leu Asp Ile Gly Val Asn Ser His Gly Lys Asn Ala Tyr Gly Glu  
 485 490 495  
 Gly Met Val Tyr Gly His Lys Tyr Phe Lys Glu Thr Asn Tyr Lys Arg  
 500 505 510  
 Leu Thr Met Val Lys Thr Arg Val Asp Pro Ser Asn Phe Phe Arg Asn  
 515 520 525  
 Glu Gln Ser Ile Pro Thr Leu Ser Ser Ser Trp Lys  
 530 535 540

(2) INFORMATION FOR SEQ ID NO: 59:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

GGTAATGATC TCCTTTCTTG TTTGACC

27

(2) INFORMATION FOR SEQ ID NO: 60:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

AGAGCGGCCG CTATATTACA ACTTCTCCAC CATCACTCCT C

41

(2) INFORMATION FOR SEQ ID NO: 61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

GGTGATGTTA ATGATAATCT CCTC

24

(2) INFORMATION FOR SEQ ID NO: 62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

AGAGCGGCCG CTACAATTCC TTCAACATGT AAATTCCTC

40

(2) INFORMATION FOR SEQ ID NO: 63:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 36 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

ACTTCCCGTA GAAACTCGGA GACTTTCACA CAATGC

36

(2) INFORMATION FOR SEQ ID NO: 64:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:

TCCATCCAAG ATCAATTCAT AAAGTGTGTC

30

(2) INFORMATION FOR SEQ ID NO: 65:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 36 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 65:

AGAGCGGCCG CTTTCATGAA CCTAGCTTCT AGTAGG

36

(2) INFORMATION FOR SEQ ID NO: 66:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 37 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:

AGAGCGGCCG CGAAATGGCC CCCCTTTTAA AACGGGG

37

(2) INFORMATION FOR SEQ ID NO: 67:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

AGAGCGGCCG CAAATGATAT CTCAGGTAA CTTTGTTCAC

40

(2) INFORMATION FOR SEQ ID NO: 68:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:

AGAGCGGCCG CATAATCAAA TAAATACACT TATGGTAACA CAG

43

(2) INFORMATION FOR SEQ ID NO: 69:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:

AGAGCGGCCG CTGGTTTTGT ATTGAGGACT CAAAACAG

38

(2) INFORMATION FOR SEQ ID NO: 70:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1757 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana

(B) STRAIN: Colombia

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: join(1..570, 801..1754)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:

ACT TCC CGT AGA AAC TCG GAG ACT TTC ACA CAA TGC CTA ACC TCA AAC	48
Thr Ser Arg Arg Asn Ser Glu Thr Phe Thr Gln Cys Leu Thr Ser Asn	
1 5 10 15	
TCC GAC CCC AAA CAT CCC ATC TCC CCC GCT ATC TTC TTC TCC GGA AAT	96
Ser Asp Pro Lys His Pro Ile Ser Pro Ala Ile Phe Phe Ser Gly Asn	
20 25 30	
GGC TCC TAC TCC TCC GTA TTA CAA GCC AAC ATC CGT AAC CTC CGC TTC	144
Gly Ser Tyr Ser Ser Val Leu Gln Ala Asn Ile Arg Asn Leu Arg Phe	
35 40 45	
AAC ACC ACC TCA ACT CCG AAA CCC TTC CTC ATA ATC GCC GCA ACA CAT	192
Asn Thr Thr Ser Thr Pro Lys Pro Phe Leu Ile Ile Ala Ala Thr His	
50 55 60	
GAA TCC CAT GTG CAA GCC GCG ATT ACT TGC GGG AAA CGC CAC AAC CTT	240
Glu Ser His Val Gln Ala Ala Ile Thr Cys Gly Lys Arg His Asn Leu	
65 70 75 80	
CAG ATG AAA ATC AGA AGT GGA GGC CAC GAC TAC GAT GGC TTG TCA TAC	288
Gln Met Lys Ile Arg Ser Gly Gly His Asp Tyr Asp Gly Leu Ser Tyr	
85 90 95	
GTT ACA TAC TCT GGC AAA CCG TTC TTC GTC CTC GAC ATG TTT AAC CTC	336
Val Thr Tyr Ser Gly Lys Pro Phe Phe Val Leu Asp Met Phe Asn Leu	
100 105 110	
CGT TCG GTG GAT GTC GAT GTG GCA AGT AAG ACC GCG TGG GTC CAA ACC	384
Arg Ser Val Asp Val Asp Val Ala Ser Lys Thr Ala Trp Val Gln Thr	
115 120 125	
GGT GCC ATA CTC GGA GAA GTT TAT TAC TAT ATA TGG GAG AAG AGC AAA	432
Gly Ala Ile Leu Gly Glu Val Tyr Tyr Tyr Ile Trp Glu Lys Ser Lys	
130 135 140	
ACC CTA GCT TAT CCC GCC GGA ATT TGT CCC ACG GTT GGT GTC GGT GGC	480
Thr Leu Ala Tyr Pro Ala Gly Ile Cys Pro Thr Val Gly Val Gly Gly	
145 150 155 160	

CAT ATC AGT GGT GGA GGT TAC GGT AAC ATG ATG AGA AAA TAC GGT CTC	528
His Ile Ser Gly Gly Gly Tyr Gly Asn Met Met Arg Lys Tyr Gly Leu	
165 170 175	
ACC GTA GAT AAT ACC ATC GAT GCA AGA ATG GTC GAC GTT AAT	570
Thr Val Asp Asn Thr Ile Asp Ala Arg Met Val Asp Val Asn	
180 185 190	
GGTATAATTG ATATCTCTAT TTTATATACT AATTAAATTT TATAGTGTGG ATCGGATAGT	630
GATTTTGGTC CATCAATTAA AAACCTGGTG AACATAAAAT TAACCAAGCA ATCAATTTAG	690
ACAAGCAACA TAATCATATA TATTTTTCTT ACATTTGTAT GTACCTGAAT ATTTATATTT	750
ATGTTTATAT GTTCTCACTA TATTTTCACT TTTGTATTTG AAAATTTTGA GGA AAA	806
Gly Lys	
ATT TTG GAT AGA AAA TTG ATG GGA GAA GAT CTC TAC TGG GCA ATA AAC	854
Ile Leu Asp Arg Lys Leu Met Gly Glu Asp Leu Tyr Trp Ala Ile Asn	
195 200 205	
GGA GGA GGA GGA GGG AGC TAC GGC GTC GTA TTG GCC TAC AAA ATA AAC	902
Gly Gly Gly Gly Gly Ser Tyr Gly Val Val Leu Ala Tyr Lys Ile Asn	
210 215 220	
CTT GTT GAA GTC CCA GAA AAC GTC ACC GTT TTC AGA ATC TCC CGG ACG	950
Leu Val Glu Val Pro Glu Asn Val Thr Val Phe Arg Ile Ser Arg Thr	
225 230 235 240	
TTA GAA CAA AAT GCG ACG GAT ATC ATT CAC CGG TGG CAA CAA GTT GCA	998
Leu Glu Gln Asn Ala Thr Asp Ile Ile His Arg Trp Gln Gln Val Ala	
245 250 255	
CCG AAG CTT CCC GAC GAG CTT TTC ATA AGA ACA GTC ATT GAC GTA GTA	1046
Pro Lys Leu Pro Asp Glu Leu Phe Ile Arg Thr Val Ile Asp Val Val	
260 265 270	
AAC GGC ACT GTT TCA TCT CAA AAG ACC GTC AGG ACA ACA TTC ATA GCA	1094
Asn Gly Thr Val Ser Ser Gln Lys Thr Val Arg Thr Thr Phe Ile Ala	
275 280 285	
ATG TTT CTA GGA GAC ACG ACA ACT CTA CTG TCG ATA TTA AAC CGG AGA	1142
Met Phe Leu Gly Asp Thr Thr Thr Leu Leu Ser Ile Leu Asn Arg Arg	
290 295 300	
TTC CCA GAA TTG GGT TTG GTC CGG TCT GAC TGT ACC GAA ACA AGC TGG	1190
Phe Pro Glu Leu Gly Leu Val Arg Ser Asp Cys Thr Glu Thr Ser Trp	
305 310 315 320	
ATC CAA TCT GTG CTA TTC TGG ACA AAT ATC CAA GTT GGT TCG TCG GAG	1238
Ile Gln Ser Val Leu Phe Trp Thr Asn Ile Gln Val Gly Ser Ser Glu	
325 330 335	
ACA CTT CTA CTC CAA AGG AAT CAA CCC GTG AAC TAC CTC AAG AGG AAA	1286
Thr Leu Leu Leu Gln Arg Asn Gln Pro Val Asn Tyr Leu Lys Arg Lys	
340 345 350	
TCA GAT TAC GTA CGT GAA CCG ATT TCA AGA ACC GGT TTA GAG TCA ATT	1334
Ser Asp Tyr Val Arg Glu Pro Ile Ser Arg Thr Gly Leu Glu Ser Ile	
355 360 365	

TGG AAG AAA ATG ATC GAG CTT GAA ATT CCG ACA ATG GCT TTC AAT CCA	1382
Trp Lys Lys Met Ile Glu Leu Glu Ile Pro Thr Met Ala Phe Asn Pro	
370 375 380	
TAC GGT GGT GAG ATG GGG AGG ATA TCA TTA CGG GTG ACT CCG TTC CCA	1430
Tyr Gly Gly Glu Met Gly Arg Ile Ser Leu Arg Val Thr Pro Phe Pro	
385 390 395 400	
TAC AGA GCC GGT AAT CTC TGG AAG ATT CAG TAC GGT GCG AAT TGG AGA	1478
Tyr Arg Ala Gly Asn Leu Trp Lys Ile Gln Tyr Gly Ala Asn Trp Arg	
405 410 415	
GAT GAG ACT TTA ACC GAC CGG TAC ATG GAA TTG ACG AGG AAG TTG TAC	1526
Asp Glu Thr Leu Thr Asp Arg Tyr Met Glu Leu Thr Arg Lys Leu Tyr	
420 425 430	
CAA TTC ATG ACA CCA TTT GTT TCC AAG AAT CCG AGA CAA TCG TTT TTC	1574
Gln Phe Met Thr Pro Phe Val Ser Lys Asn Pro Arg Gln Ser Phe Phe	
435 440 445	
AAT AAC CGT GAT GTT GAT TTG GGT ATT AAT TCT CAT AAT GGT AAA ATC	1622
Asn Asn Arg Asp Val Asp Leu Gly Ile Asn Ser His Asn Gly Lys Ile	
450 455 460	
AGT AGT TAT GTG GAA GGT AAA CGT TAC GGG AAG AAG TAT TTC GCA GGT	1670
Ser Ser Tyr Val Glu Gly Lys Arg Tyr Gly Lys Lys Tyr Phe Ala Gly	
465 470 475 480	
AAT TTC GAG AGA TTG GTG AAG ATT AAG ACG AGA GTT GAT AGT GGT AAT	1718
Asn Phe Glu Arg Leu Val Lys Ile Lys Thr Arg Val Asp Ser Gly Asn	
485 490 495	
TTC TTT AGG AAC GAA CAC AGT ATT CCT GTG TTA CCA TAA	1757
Phe Phe Arg Asn Glu His Ser Ile Pro Val Leu Pro	
500 505	

(2) INFORMATION FOR SEQ ID NO: 71:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 508 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

Thr Ser Arg Arg Asn Ser Glu Thr Phe Thr Gln Cys Leu Thr Ser Asn	
1 5 10 15	
Ser Asp Pro Lys His Pro Ile Ser Pro Ala Ile Phe Phe Ser Gly Asn	
20 25 30	
Gly Ser Tyr Ser Ser Val Leu Gln Ala Asn Ile Arg Asn Leu Arg Phe	
35 40 45	
Asn Thr Thr Ser Thr Pro Lys Pro Phe Leu Ile Ile Ala Ala Thr His	
50 55 60	

Glu Ser His Val Gln Ala Ala Ile Thr Cys Gly Lys Arg His Asn Leu  
 65 70 75 80  
 Gln Met Lys Ile Arg Ser Gly Gly His Asp Tyr Asp Gly Leu Ser Tyr  
 85 90 95  
 Val Thr Tyr Ser Gly Lys Pro Phe Phe Val Leu Asp Met Phe Asn Leu  
 100 105 110  
 Arg Ser Val Asp Val Asp Val Ala Ser Lys Thr Ala Trp Val Gln Thr  
 115 120 125  
 Gly Ala Ile Leu Gly Glu Val Tyr Tyr Tyr Ile Trp Glu Lys Ser Lys  
 130 135 140  
 Thr Leu Ala Tyr Pro Ala Gly Ile Cys Pro Thr Val Gly Val Gly Gly  
 145 150 155 160  
 His Ile Ser Gly Gly Gly Tyr Gly Asn Met Met Arg Lys Tyr Gly Leu  
 165 170 175  
 Thr Val Asp Asn Thr Ile Asp Ala Arg Met Val Asp Val Asn Gly Lys  
 180 185 190  
 Ile Leu Asp Arg Lys Leu Met Gly Glu Asp Leu Tyr Trp Ala Ile Asn  
 195 200 205  
 Gly Gly Gly Gly Gly Ser Tyr Gly Val Val Leu Ala Tyr Lys Ile Asn  
 210 215 220  
 Leu Val Glu Val Pro Glu Asn Val Thr Val Phe Arg Ile Ser Arg Thr  
 225 230 235 240  
 Leu Glu Gln Asn Ala Thr Asp Ile Ile His Arg Trp Gln Gln Val Ala  
 245 250 255  
 Pro Lys Leu Pro Asp Glu Leu Phe Ile Arg Thr Val Ile Asp Val Val  
 260 265 270  
 Asn Gly Thr Val Ser Ser Gln Lys Thr Val Arg Thr Thr Phe Ile Ala  
 275 280 285  
 Met Phe Leu Gly Asp Thr Thr Thr Leu Leu Ser Ile Leu Asn Arg Arg  
 290 295 300  
 Phe Pro Glu Leu Gly Leu Val Arg Ser Asp Cys Thr Glu Thr Ser Trp  
 305 310 315 320  
 Ile Gln Ser Val Leu Phe Trp Thr Asn Ile Gln Val Gly Ser Ser Glu  
 325 330 335  
 Thr Leu Leu Leu Gln Arg Asn Gln Pro Val Asn Tyr Leu Lys Arg Lys  
 340 345 350  
 Ser Asp Tyr Val Arg Glu Pro Ile Ser Arg Thr Gly Leu Glu Ser Ile  
 355 360 365  
 Trp Lys Lys Met Ile Glu Leu Glu Ile Pro Thr Met Ala Phe Asn Pro  
 370 375 380



Tyr Gly Gly Glu Met Gly Arg Ile Ser Leu Arg Val Thr Pro Phe Pro  
385 390 395 400

Tyr Arg Ala Gly Asn Leu Trp Lys Ile Gln Tyr Gly Ala Asn Trp Arg  
405 410 415

Asp Glu Thr Leu Thr Asp Arg Tyr Met Glu Leu Thr Arg Lys Leu Tyr  
420 425 430

Gln Phe Met Thr Pro Phe Val Ser Lys Asn Pro Arg Gln Ser Phe Phe  
435 440 445

Asn Asn Arg Asp Val Asp Leu Gly Ile Asn Ser His Asn Gly Lys Ile  
450 455 460

Ser Ser Tyr Val Glu Gly Lys Arg Tyr Gly Lys Lys Tyr Phe Ala Gly  
465 470 475 480

Asn Phe Glu Arg Leu Val Lys Ile Lys Thr Arg Val Asp Ser Gly Asn  
485 490 495

Phe Phe Arg Asn Glu His Ser Ile Pro Val Leu Pro  
500 505

(2) INFORMATION FOR SEQ ID NO: 72:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1527 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN: Colombia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1524

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

ACT TCC CGT AGA AAC TCG GAG ACT TTC ACA CAA TGC CTA ACC TCA AAC	48
Thr Ser Arg Arg Asn Ser Glu Thr Phe Thr Gln Cys Leu Thr Ser Asn	
1 5 10 15	
TCC GAC CCC AAA CAT CCC ATC TCC CCC GCT ATC TTC TTC TCC GGA AAT	96
Ser Asp Pro Lys His Pro Ile Ser Pro Ala Ile Phe Phe Ser Gly Asn	
20 25 30	
GGC TCC TAC TCC TCC GTA TTA CAA GCC AAC ATC CGT AAC CTC CGC TTC	144
Gly Ser Tyr Ser Ser Val Leu Gln Ala Asn Ile Arg Asn Leu Arg Phe	
35 40 45	

AAC ACC ACC TCA ACT CCG AAA CCC TTC CTC ATA ATC GCC GCA ACA CAT	192
Asn Thr Thr Ser Thr Pro Lys Pro Phe Leu Ile Ile Ala Ala Thr His	
50 55 60	
GAA TCC CAT GTG CAA GCC GCG ATT ACT TGC GGG AAA CGC CAC AAC CTT	240
Glu Ser His Val Gln Ala Ala Ile Thr Cys Gly Lys Arg His Asn Leu	
65 70 75 80	
CAG ATG AAA ATC AGA AGT GGA GGC CAC GAC TAC GAT GGC TTG TCA TAC	288
Gln Met Lys Ile Arg Ser Gly Gly His Asp Tyr Asp Gly Leu Ser Tyr	
85 90 95	
GTT ACA TAC TCT GGC AAA CCG TTC TTC GTC CTC GAC ATG TTT AAC CTC	336
Val Thr Tyr Ser Gly Lys Pro Phe Phe Val Leu Asp Met Phe Asn Leu	
100 105 110	
CGT TCG GTG GAT GTC GAC GTG GCA AGT AAG ACC GCG TGG GTC CAA ACC	384
Arg Ser Val Asp Val Asp Val Ala Ser Lys Thr Ala Trp Val Gln Thr	
115 120 125	
GGT GCC ATA CTC GGA GAA GTT TAT TAC TAT ATA TGG GAG AAG AGC AAA	432
Gly Ala Ile Leu Gly Glu Val Tyr Tyr Tyr Ile Trp Glu Lys Ser Lys	
130 135 140	
ACC CTA GCT TAT CCC GCC GGA ATT TGT CCC ACG GTT GGT GTC GGT GGC	480
Thr Leu Ala Tyr Pro Ala Gly Ile Cys Pro Thr Val Gly Val Gly Gly	
145 150 155 160	
CAT ATC AGT GGT GGA GGT TAC GGT AAC ATG ATG AGA AAA TAC GGT CTC	528
His Ile Ser Gly Gly Gly Tyr Gly Asn Met Met Arg Lys Tyr Gly Leu	
165 170 175	
ACC GTA GAT AAT ACC ATC GAT GCA AGA ATG GTC GAC GTA AAT GGA AAA	576
Thr Val Asp Asn Thr Ile Asp Ala Arg Met Val Asp Val Asn Gly Lys	
180 185 190	
ATT TTG GAT AGA AAA TTG ATG GGA GAA GAT CTC TAC TGG GCA ATA AAC	624
Ile Leu Asp Arg Lys Leu Met Gly Glu Asp Leu Tyr Trp Ala Ile Asn	
195 200 205	
GGA GGA GGA GGA GGG AGC TAC GGC GTC GTA TTG GCC TAC AAA ATA AAC	672
Gly Gly Gly Gly Gly Ser Tyr Gly Val Val Leu Ala Tyr Lys Ile Asn	
210 215 220	
CTT GTT GAA GTC CCA GAA AAC GTC ACC GTT TTC AGA ATC TCC CGG ACG	720
Leu Val Glu Val Pro Glu Asn Val Thr Val Phe Arg Ile Ser Arg Thr	
225 230 235 240	
TTA GAA CAA AAT GCG ACG GAT ATC ATT CAC CGG TGG CAA CAA GTT GCA	768
Leu Glu Gln Asn Ala Thr Asp Ile Ile His Arg Trp Gln Gln Val Ala	
245 250 255	
CCG AAG CTT CCC GAC GAG CTT TTC ATA AGA ACA GTC ATT GAC GTA GTA	816
Pro Lys Leu Pro Asp Glu Leu Phe Ile Arg Thr Val Ile Asp Val Val	
260 265 270	
AAC GGC ACT GTT TCA TCT CAA AAG ACC GTC AGG ACA ACA TTC ATA GCA	864
Asn Gly Thr Val Ser Ser Gln Lys Thr Val Arg Thr Thr Phe Ile Ala	
275 280 285	

ATG TTT CTA GGA GAC ACG ACA ACT CTA CTG TCG ATA TTA AAC CGG AGA	912
Met Phe Leu Gly Asp Thr Thr Thr Leu Leu Ser Ile Leu Asn Arg Arg	
290 295 300	
TTC CCA GAA TTG GGT TTG GTC CGG TCT GAC TGT ACC GAA ACA AGC TGG	960
Phe Pro Glu Leu Gly Leu Val Arg Ser Asp Cys Thr Glu Thr Ser Trp	
305 310 315 320	
ATC CAA TCT GTG CTA TTC TGG ACA AAT ATC CAA GTT GGT TCG TCG GAG	1008
Ile Gln Ser Val Leu Phe Trp Thr Asn Ile Gln Val Gly Ser Ser Glu	
325 330 335	
ACA CTT CTA CTC CAA AGG AAT CAA CCC GTG AAC TAC CTC AAG AGG AAA	1056
Thr Leu Leu Leu Gln Arg Asn Gln Pro Val Asn Tyr Leu Lys Arg Lys	
340 345 350	
TCA GAT TAC GTA CGT GAA CCG ATT TCA AGA ACC GGT TTA GAG TCA ATT	1104
Ser Asp Tyr Val Arg Glu Pro Ile Ser Arg Thr Gly Leu Glu Ser Ile	
355 360 365	
TGG AAG AAA ATG ATC GAG CTT GAA ATT CCG ACA ATG GCT TTC AAT CCA	1152
Trp Lys Lys Met Ile Glu Leu Glu Ile Pro Thr Met Ala Phe Asn Pro	
370 375 380	
TAC GGT GGT GAG ATG GGG AGG ATA TCA TCT ACG GTG ACT CCG TTC CCA	1200
Tyr Gly Gly Glu Met Gly Arg Ile Ser Ser Thr Val Thr Pro Phe Pro	
385 390 395 400	
TAC AGA GCC GGT AAT CTC TGG AAG ATT CAG TAC GGT GCG AAT TGG AGA	1248
Tyr Arg Ala Gly Asn Leu Trp Lys Ile Gln Tyr Gly Ala Asn Trp Arg	
405 410 415	
GAT GAG ACT TTA ACC GAC CGG TAC ATG GAA TTG ACG AGG AAG TTG TAC	1296
Asp Glu Thr Leu Thr Asp Arg Tyr Met Glu Leu Thr Arg Lys Leu Tyr	
420 425 430	
CAA TTC ATG ACA CCA TTT GTT TCC AAG AAT CCG AGA CAA TCG TTT TTC	1344
Gln Phe Met Thr Pro Phe Val Ser Lys Asn Pro Arg Gln Ser Phe Phe	
435 440 445	
AAT TAC CGT GAT GTT GAT TTG GGT ATT AAT TCT CAT AAT GGT AAA ATC	1392
Asn Tyr Arg Asp Val Asp Leu Gly Ile Asn Ser His Asn Gly Lys Ile	
450 455 460	
AGT AGT TAT GTG GAA GGT AAA CGT TAC GGG AAG AAG TAT TTC GCA GGT	1440
Ser Ser Tyr Val Glu Gly Lys Arg Tyr Gly Lys Lys Tyr Phe Ala Gly	
465 470 475 480	
AAT TTC GAG AGA TTG GTG AAG ATT AAG ACG AGA GTT GAT AGT GGT AAT	1488
Asn Phe Glu Arg Leu Val Lys Ile Lys Thr Arg Val Asp Ser Gly Asn	
485 490 495	
TTC TTT AGG AAC GAA CAG AGT ATT CCT GTG TTA CCA TAA	1527
Phe Phe Arg Asn Glu Gln Ser Ile Pro Val Leu Pro	
500 505	

(2) INFORMATION FOR SEQ ID NO: 73:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 508 amino acids

(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

Thr Ser Arg Arg Asn Ser Glu Thr Phe Thr Gln Cys Leu Thr Ser Asn  
1 5 10 15  
Ser Asp Pro Lys His Pro Ile Ser Pro Ala Ile Phe Phe Ser Gly Asn  
20 25 30  
Gly Ser Tyr Ser Ser Val Leu Gln Ala Asn Ile Arg Asn Leu Arg Phe  
35 40 45  
Asn Thr Thr Ser Thr Pro Lys Pro Phe Leu Ile Ile Ala Ala Thr His  
50 55 60  
Glu Ser His Val Gln Ala Ala Ile Thr Cys Gly Lys Arg His Asn Leu  
65 70 75 80  
Gln Met Lys Ile Arg Ser Gly Gly His Asp Tyr Asp Gly Leu Ser Tyr  
85 90 95  
Val Thr Tyr Ser Gly Lys Pro Phe Phe Val Leu Asp Met Phe Asn Leu  
100 105 110  
Arg Ser Val Asp Val Asp Val Ala Ser Lys Thr Ala Trp Val Gln Thr  
115 120 125  
Gly Ala Ile Leu Gly Glu Val Tyr Tyr Tyr Ile Trp Glu Lys Ser Lys  
130 135 140  
Thr Leu Ala Tyr Pro Ala Gly Ile Cys Pro Thr Val Gly Val Gly Gly  
145 150 155 160  
His Ile Ser Gly Gly Gly Tyr Gly Asn Met Met Arg Lys Tyr Gly Leu  
165 170 175  
Thr Val Asp Asn Thr Ile Asp Ala Arg Met Val Asp Val Asn Gly Lys  
180 185 190  
Ile Leu Asp Arg Lys Leu Met Gly Glu Asp Leu Tyr Trp Ala Ile Asn  
195 200 205  
Gly Gly Gly Gly Gly Ser Tyr Gly Val Val Leu Ala Tyr Lys Ile Asn  
210 215 220  
Leu Val Glu Val Pro Glu Asn Val Thr Val Phe Arg Ile Ser Arg Thr  
225 230 235 240  
Leu Glu Gln Asn Ala Thr Asp Ile Ile His Arg Trp Gln Gln Val Ala  
245 250 255  
Pro Lys Leu Pro Asp Glu Leu Phe Ile Arg Thr Val Ile Asp Val Val  
260 265 270  
Asn Gly Thr Val Ser Ser Gln Lys Thr Val Arg Thr Thr Phe Ile Ala  
275 280 285

Met Phe Leu Gly Asp Thr Thr Thr Leu Leu Ser Ile Leu Asn Arg Arg  
 290 295 300  
 Phe Pro Glu Leu Gly Leu Val Arg Ser Asp Cys Thr Glu Thr Ser Trp  
 305 310 315 320  
 Ile Gln Ser Val Leu Phe Trp Thr Asn Ile Gln Val Gly Ser Ser Glu  
 325 330 335  
 Thr Leu Leu Leu Gln Arg Asn Gln Pro Val Asn Tyr Leu Lys Arg Lys  
 340 345 350  
 Ser Asp Tyr Val Arg Glu Pro Ile Ser Arg Thr Gly Leu Glu Ser Ile  
 355 360 365  
 Trp Lys Lys Met Ile Glu Leu Glu Ile Pro Thr Met Ala Phe Asn Pro  
 370 375 380  
 Tyr Gly Gly Glu Met Gly Arg Ile Ser Ser Thr Val Thr Pro Phe Pro  
 385 390 395 400  
 Tyr Arg Ala Gly Asn Leu Trp Lys Ile Gln Tyr Gly Ala Asn Trp Arg  
 405 410 415  
 Asp Glu Thr Leu Thr Asp Arg Tyr Met Glu Leu Thr Arg Lys Leu Tyr  
 420 425 430  
 Gln Phe Met Thr Pro Phe Val Ser Lys Asn Pro Arg Gln Ser Phe Phe  
 435 440 445  
 Asn Tyr Arg Asp Val Asp Leu Gly Ile Asn Ser His Asn Gly Lys Ile  
 450 455 460  
 Ser Ser Tyr Val Glu Gly Lys Arg Tyr Gly Lys Lys Tyr Phe Ala Gly  
 465 470 475 480  
 Asn Phe Glu Arg Leu Val Lys Ile Lys Thr Arg Val Asp Ser Gly Asn  
 485 490 495  
 Phe Phe Arg Asn Glu Gln Ser Ile Pro Val Leu Pro  
 500 505

(2) INFORMATION FOR SEQ ID NO: 74:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1530 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: Colombia

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..1527

• (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:

TCC ATC CAA GAT CAA TTC ATA AAC TGT GTC AAA AGA AAC ACA CAT GTT	48
Ser Ile Gln Asp Gln Phe Ile Asn Cys Val Lys Arg Asn Thr His Val	
1 5 10 15	
TCT TTT CCA CTC GAG AAA ACG TTA TTC ACC CCT GCG AAA AAC GTC TCT	96
Ser Phe Pro Leu Glu Lys Thr Leu Phe Thr Pro Ala Lys Asn Val Ser	
20 25 30	
TTG TTC AAC CAA GTC CTT GAA TCG ACG GCT CAA AAT CTC CAG TTC TTG	144
Leu Phe Asn Gln Val Leu Glu Ser Thr Ala Gln Asn Leu Gln Phe Leu	
35 40 45	
GCA AAA TCC ATG CCT AAA CCG GGA TTC ATA TTC AGA CCG ATT CAC CAG	192
Ala Lys Ser Met Pro Lys Pro Gly Phe Ile Phe Arg Pro Ile His Gln	
50 55 60	
TCT CAA GTC CAA GCT TCC ATC ATT TGT TCA AAG AAA CTC GGA ATT CAT	240
Ser Gln Val Gln Ala Ser Ile Ile Cys Ser Lys Lys Leu Gly Ile His	
65 70 75 80	
TTT CGT GTT AGA AGT GGC GGT CAC GAT TTC GAG GCC TTG TCT TAT GTT	288
Phe Arg Val Arg Ser Gly Gly His Asp Phe Glu Ala Leu Ser Tyr Val	
85 90 95	
TCA CGG ATT GAA AAA CCG TTT ATA TTA CTC GAC CTG TCA AAA TTG AAA	336
Ser Arg Ile Glu Lys Pro Phe Ile Leu Leu Asp Leu Ser Lys Leu Lys	
100 105 110	
CAA ATC AAT GTT GAT ATT GAA TCC AAT AGT GCT TGG GTT CAA CCT GGT	384
Gln Ile Asn Val Asp Ile Glu Ser Asn Ser Ala Trp Val Gln Pro Gly	
115 120 125	
GCT ACG CTT GGT GAG CTT TAC TAC AGA ATT GCA GAG AAG AGC AAG ATC	432
Ala Thr Leu Gly Glu Leu Tyr Tyr Arg Ile Ala Glu Lys Ser Lys Ile	
130 135 140	
CAT GGA TTT CCC GCG GGT TTG TGC ACA AGT GTA GGC ATA GGT GGG TAT	480
His Gly Phe Pro Ala Gly Leu Cys Thr Ser Val Gly Ile Gly Gly Tyr	
145 150 155 160	
ATG ACA GGC GGT GGA TAC GGT ACC TTG ATG AGG AAG TAT GGT CTT GCG	528
Met Thr Gly Gly Gly Tyr Gly Thr Leu Met Arg Lys Tyr Gly Leu Ala	
165 170 175	
GGA GAT AAT GTT CTA GAC GTA AAG ATG GTT GAT GCA AAT GGT AAA TTA	576
Gly Asp Asn Val Leu Asp Val Lys Met Val Asp Ala Asn Gly Lys Leu	
180 185 190	
CTC GAC AGA GCC GCG ATG GGT GAG GAC CTA TTT TGG GCG ATT AGA GGA	624
Leu Asp Arg Ala Ala Met Gly Glu Asp Leu Phe Trp Ala Ile Arg Gly	
195 200 205	
GGC GGT GGA GCG AGT TTC GGG ATA GTT CTA GCA TGG AAG ATC AAG CTT	672
Gly Gly Gly Ala Ser Phe Gly Ile Val Leu Ala Trp Lys Ile Lys Leu	
210 215 220	

GTT CCT GTT CCT AAG ACT GTT ACC GTC TTC ACT GTC ACC AAA ACG TTA	720
Val Pro Val Pro Lys Thr Val Thr Val Phe Thr Val Thr Lys Thr Leu	
225 230 235 240	
GAA CAA GAC GCA AGA TTG AAG ACT ATT TCT AAG TGG CAA CAA ATT TCA	768
Glu Gln Asp Ala Arg Leu Lys Thr Ile Ser Lys Trp Gln Gln Ile Ser	
245 250 255	
TCC AAG ATT ATT GAA GAG ATA CAC ATC CGA GTG GTA CTC AGA GCA GCT	816
Ser Lys Ile Ile Glu Glu Ile His Ile Arg Val Val Leu Arg Ala Ala	
260 265 270	
GGA AAT GAT GGA AAC AAG ACT GTG ACA ATG ACC TAC CTA GGT CAG TTT	864
Gly Asn Asp Gly Asn Lys Thr Val Thr Met Thr Tyr Leu Gly Gln Phe	
275 280 285	
CTT GGC GAG AAA GGC ACC TTG CTG AAG GTT ATG GAG AAG GCT TTT CCA	912
Leu Gly Glu Lys Gly Thr Leu Leu Lys Val Met Glu Lys Ala Phe Pro	
290 295 300	
GAA CTA GGG TTA ACT CAA AAG GAT TGT ACT GAA ATG AGC TGG ATT GAA	960
Glu Leu Gly Leu Thr Gln Lys Asp Cys Thr Glu Met Ser Trp Ile Glu	
305 310 315 320	
GCC GCC CTT TTC CAT GGT GGA TTT CCA ACA GGT TCT CCT ATT GAA ATT	1008
Ala Ala Leu Phe His Gly Gly Phe Pro Thr Gly Ser Pro Ile Glu Ile	
325 330 335	
TTG CTT CAG CTC AAG TCG CCT CTA GGA AAA GAT TAC TTC AAA GCA ACG	1056
Leu Leu Gln Leu Lys Ser Pro Leu Gly Lys Asp Tyr Phe Lys Ala Thr	
340 345 350	
TCG GAT TTC GTT AAA GAA CCT ATT CCT GTG ATA GGC TTC AAA GGA ATA	1104
Ser Asp Phe Val Lys Glu Pro Ile Pro Val Ile Gly Phe Lys Gly Ile	
355 360 365	
TTC AAA AGA TTG ATT GAA GGA AAC ACA ACA TTT CTG AAC TGG ACT CCT	1152
Phe Lys Arg Leu Ile Glu Gly Asn Thr Thr Phe Leu Asn Trp Thr Pro	
370 375 380	
TAC GGT GGT ATG ATG TCG AAA ATC CCT GAA TCT GCG ATC CCA TTT CCG	1200
Tyr Gly Gly Met Met Ser Lys Ile Pro Glu Ser Ala Ile Pro Phe Pro	
385 390 395 400	
CAT AGA AAC GGA ACC CTC TTC AAG ATT CTC TAT TAC GCG AAC TGG CTA	1248
His Arg Asn Gly Thr Leu Phe Lys Ile Leu Tyr Tyr Ala Asn Trp Leu	
405 410 415	
GAG AAT GAC AAG ACA TCG AGT AGA AAA ATC AAC TGG ATC AAA GAG ATA	1296
Glu Asn Asp Lys Thr Ser Ser Arg Lys Ile Asn Trp Ile Lys Glu Ile	
420 425 430	
TAC AAT TAC ATG GCG CCT TAT GTC TCA AGC AAT CCA AGA CAA GCA TAT	1344
Tyr Asn Tyr Met Ala Pro Tyr Val Ser Ser Asn Pro Arg Gln Ala Tyr	
435 440 445	
GTG AAC TAC AGA GAT CTA GAC TTC GGA CAG AAC AAG AAC AAC GCA AAG	1392
Val Asn Tyr Arg Asp Leu Asp Phe Gly Gln Asn Lys Asn Asn Ala Lys	
450 455 460	
GTT AAC TTC ATT GAA GCT AAA ATC TGG GGA CCT AAG TAC TTC AAA GGC	1440

Val Asn Phe Ile Glu Ala Lys Ile Trp Gly Pro Lys Tyr Phe Lys Gly  
 465 470 475 480

AAT TTT GAC AGA TTG GTG AAG ATT AAA ACC AAG GTT GAT CCA GAG AAC 1488  
 Asn Phe Asp Arg Leu Val Lys Ile Lys Thr Lys Val Asp Pro Glu Asn  
 485 490 495

TTC TTC AGG CAC GAG CAG AGT ATC CCA CCT ATG CCC TAC TAG 1530  
 Phe Phe Arg His Glu Gln Ser Ile Pro Pro Met Pro Tyr  
 500 505

(2) INFORMATION FOR SEQ ID NO: 75:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 509 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

Ser Ile Gln Asp Gln Phe Ile Asn Cys Val Lys Arg Asn Thr His Val  
 1 5 10 15

Ser Phe Pro Leu Glu Lys Thr Leu Phe Thr Pro Ala Lys Asn Val Ser  
 20 25 30

Leu Phe Asn Gln Val Leu Glu Ser Thr Ala Gln Asn Leu Gln Phe Leu  
 35 40 45

Ala Lys Ser Met Pro Lys Pro Gly Phe Ile Phe Arg Pro Ile His Gln  
 50 55 60

Ser Gln Val Gln Ala Ser Ile Ile Cys Ser Lys Lys Leu Gly Ile His  
 65 70 75 80

Phe Arg Val Arg Ser Gly Gly His Asp Phe Glu Ala Leu Ser Tyr Val  
 85 90 95

Ser Arg Ile Glu Lys Pro Phe Ile Leu Leu Asp Leu Ser Lys Leu Lys  
 100 105 110

Gln Ile Asn Val Asp Ile Glu Ser Asn Ser Ala Trp Val Gln Pro Gly  
 115 120 125

Ala Thr Leu Gly Glu Leu Tyr Tyr Arg Ile Ala Glu Lys Ser Lys Ile  
 130 135 140

His Gly Phe Pro Ala Gly Leu Cys Thr Ser Val Gly Ile Gly Gly Tyr  
 145 150 155 160

Met Thr Gly Gly Gly Tyr Gly Thr Leu Met Arg Lys Tyr Gly Leu Ala  
 165 170 175

Gly Asp Asn Val Leu Asp Val Lys Met Val Asp Ala Asn Gly Lys Leu  
 180 185 190

Leu Asp Arg Ala Ala Met Gly Glu Asp Leu Phe Trp Ala Ile Arg Gly  
 195 200 205



Gly Gly Gly Ala Ser Phe Gly Ile Val Leu Ala Trp Lys Ile Lys Leu  
210 215 220

Val Pro Val Pro Lys Thr Val Thr Val Phe Thr Val Thr Lys Thr Leu  
225 230 235 240

Glu Gln Asp Ala Arg Leu Lys Thr Ile Ser Lys Trp Gln Gln Ile Ser  
245 250 255

Ser Lys Ile Ile Glu Glu Ile His Ile Arg Val Val Leu Arg Ala Ala  
260 265 270

Gly Asn Asp Gly Asn Lys Thr Val Thr Met Thr Tyr Leu Gly Gln Phe  
275 280 285

Leu Gly Glu Lys Gly Thr Leu Leu Lys Val Met Glu Lys Ala Phe Pro  
290 295 300

Glu Leu Gly Leu Thr Gln Lys Asp Cys Thr Glu Met Ser Trp Ile Glu  
305 310 315 320

Ala Ala Leu Phe His Gly Gly Phe Pro Thr Gly Ser Pro Ile Glu Ile  
325 330 335

Leu Leu Gln Leu Lys Ser Pro Leu Gly Lys Asp Tyr Phe Lys Ala Thr  
340 345 350

Ser Asp Phe Val Lys Glu Pro Ile Pro Val Ile Gly Phe Lys Gly Ile  
355 360 365

Phe Lys Arg Leu Ile Glu Gly Asn Thr Thr Phe Leu Asn Trp Thr Pro  
370 375 380

Tyr Gly Gly Met Met Ser Lys Ile Pro Glu Ser Ala Ile Pro Phe Pro  
385 390 395 400

His Arg Asn Gly Thr Leu Phe Lys Ile Leu Tyr Tyr Ala Asn Trp Leu  
405 410 415

Glu Asn Asp Lys Thr Ser Ser Arg Lys Ile Asn Trp Ile Lys Glu Ile  
420 425 430

Tyr Asn Tyr Met Ala Pro Tyr Val Ser Ser Asn Pro Arg Gln Ala Tyr  
435 440 445

Val Asn Tyr Arg Asp Leu Asp Phe Gly Gln Asn Lys Asn Asn Ala Lys  
450 455 460

Val Asn Phe Ile Glu Ala Lys Ile Trp Gly Pro Lys Tyr Phe Lys Gly  
465 470 475 480

Asn Phe Asp Arg Leu Val Lys Ile Lys Thr Lys Val Asp Pro Glu Asn  
485 490 495

Phe Phe Arg His Glu Gln Ser Ile Pro Pro Met Pro Tyr  
500 505